Toward Simulation of Complex Reactive Systems: Development and Application of Enhanced Sampling Methods

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Evgenii Fetisov

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

J. Ilja Siepmann

March, 2018

Acknowledgements

There are many people who deserve my gratitude and to whom I became indebted during my time at the graduate school. Above all, I would like to thank my advisor, Prof. J. Ilja Siepmann, who introduced me to the exciting world of statistical mechanics and molecular simulations. His guidance and high requirements for scientific integrity helped me to become a better scientist and a better problem-solver.

I would also want to thank numerous collaborators that drove my research projects forward, but namely, Will Kuo at Lawrence Livermore National Laboratory, Chris Mundy at Pacific Northwest National Laboratory, and Chris Knight at Argonne National Laboratory. Additionally, their mentorship helped me a lot in understanding many sides of research environment.

I also thank all past and present members of the Siepmann group with whom I had a chance to interact for their friendship and fruitful discussions. Furthermore, I thank my significant other for her eternal patience and her moral support when it was needed. Likewise, I thank my family for always supporting my life choices and career directions.

This work would not be possible without financial support from National Science Foundation and Department of Energy, as well as from University of Minnesota through Graduate School Doctoral Dissertation Fellowship, and without computer time provided by Minnesota Supercomputing Institute and Argonne Leadership Computing Facility.

Abstract

Predictive modeling of fluid phase and sorption equilibria for reacting systems presents one of the grand challenges in the field of molecular simulation. Difficulties in the study of such systems arise from the need (i) to accurately model both strong, short-ranged interactions leading to the formation of chemical bonds and weak interactions representing the environment, and (ii) to sample the range of time scales involving frequent molecular collisions, slow diffusion, and infrequent reactive events. This thesis showcases some of my efforts in developing and applying advanced simulation methods to a variety of important systems.

Chapters 2 and 3 describe how a novel Monte Carlo method (reactive first principles Monte Carlo or RxFPMC) can be used to overcome some limitations of existing methods for simulation of reactive systems. Chapter 4 shows how advanced sampling techniques in combination with sophisticated interatomic potentials can be used to elucidate nucleation pathways. Chapters 5 and 6 manifest how first principles simulations can be leveraged to understand liquid structure of novel complex solvents as well as reactive processes in such solvents. Finally, the last chapter discusses the use of smart sampling algorithms to study chemisorption of mixed ligands on nanoparticles.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The main purpose of molecular simulations is to provide macroscopic properties of a system based on the set of microscopic (atomic) configurations. Such connection between macroscopic and microscopic representations is solely built upon and rigorously derived from the principles of statistical mechanics [1, 2]. Additionally, molecular simulations can provide important physical insights into systems' behavior that are not accessible in real-world experiments or even provide guidance for new investigations. So it is not surprising that computer experiments became standard in chemists' and engineers' toolbox over the past two decades.

Two major methods of generating an appropriate set (ensemble) of atomic configurations are molecular dynamics (MD) and Monte Carlo (MC) approaches. These two methods in their basic forms have been extensively reviewed in various excellent books [3, 4] and will not be discussed here. The goal of this thesis, however, is to show how advanced sampling techniques in combination with complex interatomic potentials (e.g., derived from density functional theory) can improve our existing simulation capabilities and yield molecular-level understanding of important technological systems. Before proceeding to the results, a quick overview of some existing techniques for advanced sampling of molecular phase space is provided.

1.1 Enhanced Sampling Techniques

1.1.1 Reaction Ensemble Monte Carlo Method

One of the main advantages of the MC approach over MD is the availability of unphysical moves that can directly overcome kinetic barriers that hinder the efficiency of MD simulations. A prime example of kinetically-limited processes is reactive events in bulk systems, where there can be substantial activation barriers to overcome. MC framework provides a mean to obtain equilibrium distribution of reactants and products via introduction of smart MC moves that artificially overcome the barriers. The first Monte Carlo algorithm to study chemical equilibria in reactive systems was introduced in early 1990's by Shaw [5], and three years later, a new algorithm was independently developed by Smith and Triska [6] and Johnson et al. [7] and had been named Reaction Ensemble Monte Carlo (RxMC) method. The method is applicable to calculations at either fixed (T, V) or (p, T), and to systems undergoing any number of simultaneous reactions, which occur in any distribution of phases.

In the RxMC approach two pieces of information have to be specified in advance. The first is the set of possible reactions. It should be noted that RxMC can only predict equilibrium compositions of the specified reactions and new species cannot be created unless they have been specified in the reaction set. Another piece of information includes the ideal gas partition function, which includes intramolecular (vibrational, rotational, and electronic) contributions for both reactants and products.

Any chemical reaction occurring in one phase at fixed (T, V) can be expressed as

$$\sum_{i=1}^{s} \nu_i S_i = 0, \tag{1.1}$$

where, ν_i is the stoichiometric number of species i (which can be positive or negative), S_i is the chemical symbol of species i, and s is the total number of chemical species participating in the reaction. In order to simulate this transformation, a special type of MC move is introduced. This move consists of the simultaneous insertion and deletion of molecules in the system under study according to the reaction stoichometry with appropriate acceptance probabilities to ensure microscopic reversibility. This guarantees

the reaction equilibrium to be established

$$\sum_{i=1}^{s} \nu_i \mu_i = 0, \tag{1.2}$$

where μ_i is the chemical potential of species i. Changes in the number of particles ($N_i \geq 0$) due to the reaction MC steps must satisfy the law of conservation of mass

$$N_i = N_i^0 + \nu_i \xi, \tag{1.3}$$

where ξ is the so-called extent of reaction and N_i^0 is the number of particles of species i prior to the reaction step. In the RxMC algorithm, ξ is an integer and is usually equal to $+1$ for a forward step and to -1 for a reverse step. The general form of acceptance probabilities used in RxMC is then:

$$P_{kl}^{\xi} = \min \left\{ \left(1, \left(\beta p^0 V\right)^{\bar{\nu}\xi} K^{\xi} \prod_{i=1}^{s} \left[\frac{(N_i^0)!}{(N_i^0 + \nu_i \xi)!} \right] \exp(-\beta \Delta U_{kl})\right) \right\}, \tag{1.4}$$

where $\beta = 1/k_{\mathrm{B}}T$, with k_{B} being the Boltzmann constant and T being absolute temperature, p^0 is standard pressure, V is volume of the system, $\bar{\nu}$ is the total change in the number of molecules, K is ideal gas equilibrium constant, and ΔU_{kl} is the change in internal energy. In turn, the overall RxMC approach is composed of several types of MC steps:

- Particle displacements/reorientations;
- Reactions:
    1. A reaction is randomly selected if there are various reactions in the reaction set;
    2. A direction of the reaction is randomly selected, forward or reverse;
    3. A set of reactants and products in the system is randomly selected according to the chosen reaction stoichometry;
    4. The reaction move is attempted according to the acceptance probability provided above;

- Volume changes in case of constant P simulations.

1.1.2 Aggregation-Volume-Bias Monte Carlo

Systems consisting of strongly associating molecules are an important class of systems presenting various interesting physical and chemical properties. Due to formation of aggregates in such systems, they show non-ideal behavior even in low-density gas phases and low-concentration solutions. Usually, this type of association is due to specific intermolecular interactions with deep but narrow wells (e.g., hydrogen bonds). Typical examples of associating fluids are water, alcohols and organic acids. Additionally, molecules in reactive systems can be considered as collections of associating atoms. Molecular simulations of strongly associating systems present a challenging task because a given system under investigation can be "trapped" in extremely stable bonded configurations. The conventional Metropolis MC algorithm [8] is not able to sample the phase space of such systems in an efficient manner either because of these trapped configurations, which have large Boltzmann weight, or because there is a relatively low probability of finding these configurations using a symmetric underlying transition matrix in the Markov chain. Similar limitations apply to MD simulations of strongly associating systems. Therefore, an efficient algorithm is needed able to overcome these difficulties.

Over the last two decades, various biased MC algorithms, which differ in generality, complexity, and applicability have been proposed for this purpose. Some examples include association-biased Monte Carlo (ABMC) method [9], bond-bias Monte Carlo (BBMC) method [10], and monomer-addition-substraction algorithm (MASA) [11]. Besides these three methods, the relatively simpler and more general aggregation-volume-bias Monte Carlo (AVBMC) method was proposed by Chen and Siepmann [12, 13] to overcome some of the drawbacks of the previously existing algorithms. For example, the ABMC method [14] involves two-particle displacements and requires specific determination of bonding regions, which dramatically increases computational costs and complexity. The BBMC [10] method is limited to systems, where at most one other molecule may be in another molecule's bonding region and where this bonding region is exactly known. The MASA algorithm [11] is limited only to bivalent associating fluids (e.g., hydrogen fluoride), since it prevents the formation of branched aggregates. The

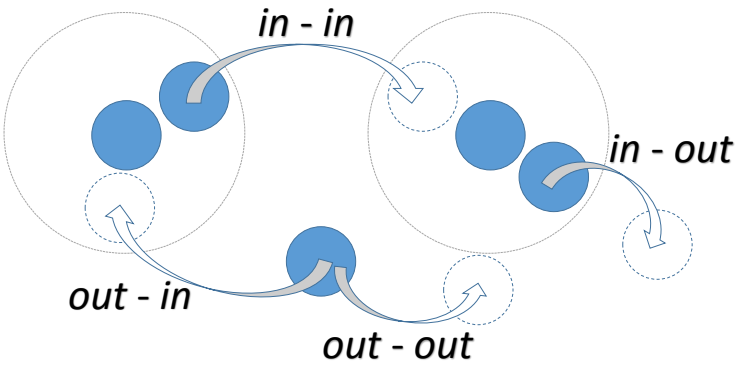

Figure 1.1: Schematic representation of the AVBMC moves.

AVBMC algorithm, especially in its improved form, overcomes these drawbacks, but at the same time retains robustness and generality.

To understand how such limitations can be lifted through design of "smart" MC moves, a discussion of ideas behind AVBMC is presented. As mentioned above, the conventional Metropolis algorithm results in a low efficiency of locating the bonded configurations of a given system. To overcome this problem, the AVBMC algorithm introduces a biased intrabox (interbox moves can also be used) swap move, which allows for efficient sampling of aggregates by allowing direct hopping between the bonded and non-bonded configurations, where bonded configurations are defined by a radius around a bonding site. This is done by increasing the transition probability for moves that lead to the formation of bonded configurations and enhancing the acceptance rate for moves that lead to their destruction. Schematic representation of all possible AVBMC moves is given in Figure 1.1.

An AVBMC move proceeds as follows:

1. A molecule i to be swapped is selected randomly;

2. A second molecule j acting as the target for the move is randomly selected;

3. With a probability P_{bias}, molecule i is only allowed to swap into bonded region of molecule j, called the B_{in} state, whereas with a probability of $1 - P_{\text{bias}}$, the molecule i is swapped into the nonbonded region of molecule j, called the B_{out}

state;

4. The potential energy difference, ΔE is calculated;

5. The move is accepted according to the following set of acceptance probabilities:

   - in the cases when the move does not involve molecule i entering or leaving the bonded region of particle j, that is, in the cases $in \to in$ or $out \to out$, the standard Metropolis acceptance rule [15] is used;
   - in the case in which the move results in the formation of a bonded configuration of the molecules i and j, that is in the case $out \to in$, the following acceptance rule is used:

$$P_{\text{acc}}(A_{\text{out}} \to B_{\text{in}}) = \min\left\{1, \frac{(1-P_{\text{bias}})V_{\text{in}}\exp(-\Delta E/k_{\text{B}}T)}{P_{\text{bias}}V_{\text{out}}}\right\}, \tag{1.5}$$

     where V_{in} is the volume of the bonded region and $V_{\text{out}} = V - V_{\text{in}}$ is the remainder of the system's volume;

   - in the case in which the move results in the destruction of a bonded configuration of the molecules i and j, that is in the case $in \to out$, the following acceptance rule is used:

$$P_{\text{acc}}(A_{\text{in}} \to B_{\text{out}}) = \min\left\{1, \frac{P_{\text{bias}}V_{\text{out}}\exp(-\Delta E/k_{\text{B}}T)}{(1-P_{\text{bias}})V_{\text{in}}}\right\}. \tag{1.6}$$

One can see from the algorithm provided above that underlying matrix of the Markov chain is asymmetric, and, consequently, to ensure the detailed balance condition, the acceptance rules were written to introduce bias proportional to $P_{\text{bias}}/V_{\text{in}}$ for trials that are placed in the *in* region and proportional to $(1-P_{\text{bias}})/V_{\text{out}}$ for trials that are placed in the *out* region.

However, It was pointed out by Wierzchowski and Kofke [16] that the original AVBMC scheme (AVBMC1) had some shortcomings. First, $in \to in$ and $out \to out$ moves do not help much with sampling compared to the conventional swap moves, and, secondly, the majority of AVBMC moves belong to $out \to out$ or $out \to in$ types due to the low probability of finding molecule i in the bonded region of a molecule j. To overcome these drawbacks, two modified AVBMC schemes (AVBMC2 and AVBMC3) were

proposed [13]. The AVBMC2 scheme simply restricts all swap moves to either $in \rightarrow out$ or $out \rightarrow in$, therefore leading to a more frequent sampling of cluster formations and destructions.

The AVBMC3 scheme is similar to multiphase Gibbs ensemble technique [17, 18] and can be viewed as an extension of the AVBMC2 method. It allows frequent sampling of cluster hopping through the introduction of $in \rightarrow in$ moves between different aggregates. This method is usually used to study systems where clusters of different sizes can be found and where concentration of free monomers might be small. It should be noted that in AVBMC2 and AVBMC3 algorithms, as compared to the original AVBMC1 algorithm, one needs to keep track of the number of neighbors of each molecule, which increases computational cost of the simulations. However, this associated cost is not large compared to efficiency gained and the improved algorithms should be used instead of the original one. Overall, AVBMC algorithms have a lot of advantages over conventional Metropolis MC scheme when it comes to more efficient sampling of the relevant microphase regions in strongly associating fluids. Most notably, they have been used to study several vapor-to-liquid nucleation phenomena [19, 20], but in general, these algorithms are applicable to any systems exhibiting cluster formation, including reactive systems.

1.1.3 Umbrella Sampling

Sampling phase space regions with small Boltzmann factors, $\exp(-\beta\mathcal{H})$ ($\mathcal{H}$ is the Hamiltonian of a system), using conventional MD or MC simulations requires exceedingly long simulation times in order to converge to statistically meaningful results, especially, when accurate free energy differences are desired. Several methods for free energy calculations have been introduced over the years with the goal of overcoming the issue. Some examples include, but are not limited to, metadynamics [21], free energy perturbation theory [22], Wang–Landau sampling [23], and thermodynamic integration [24]. Additionally, Torrie and Valleau developed umbrella sampling technique, where an external biasing potential is used, which drives simulated systems into regions of low probability [25]. Subsequently, the biasing potential can be removed and the unbiased free energy of a system can be back-calculated.

Here, we briefly discuss the idea behind umbrella sampling since it is a relatively

simple, but at the same time, very powerful method of computing relative free energy differences. First, the biased free energy profile, $F_{\mathrm{b}}(q)$ is obtained by adding an external potential, $U_{\mathrm{b}}(q)$ (usually in the form of a harmonic potential, $U_{\mathrm{b}}(q) = \frac{k}{2}(q - q_0)^2$):

$$F_{\mathrm{b}}(q) = F(q) + U_{\mathrm{b}}(q), \tag{1.7}$$

where $F(q)$ is the true free energy of a system for a specific value of a collective variable q. Then, $F(q)$ is reconstructed by an unbiasing procedure. This simple approach is often used in nucleation simulations, where one is interested in free energies of nucleation for one given cluster [19]. However in practice, to ensure sufficient overlap between successive states, a series of biasing potentials $\{U_{\mathrm{b}}(q_i)\}$ along the reaction coordinate q is introduced instead and separate MD or MC simulations are performed in each "window", which essentially dives the system along the reaction coordinate. The unbiasing procedure in this case is known as the weighted histogram analysis method (WHAM) [26]. It involves connecting the resulting probability distributions together through a self-consistent procedure seeking to minimize the total statistical error in the resulting distribution. The main assumption behind the stitching procedure is that the quality of sampling in each in each window is approximately the same, which is usually true for sufficiently long simulations. Dividing a given reaction coordinate into several regions is also beneficial from the standpoint of parallel efficiency: all individual windows can be run at the same time given that enough computational resources are available.

1.1.4 Molecule Identity Exchange

Monte Carlo approach offers a great flexibility when it comes to design of trial moves. One of the examples of such sophisticated moves are identity exchange moves. Initially, they were designed to sample spatial distribution of molecules of different lengths in monolayers [27]. A typical move consists of random selection of two molecules with varying chain lengths and then simultaneous exchange of terminal group while the backbone remains at the same location. As an example, consider an exchange between hexane and dodecane molecules: in this case, pentyl unit will remain in place while methyl unit of the hexane molecule is replaced with heptyl and heptyl unit of the dodecane molecule is

replaced with methyl. To further increase efficiency of such moves, configurational-bias Monte Carlo method [28, 29] is used to grow the exchanged units. This combined identity exchange moves greatly increases sampling efficiency, especially, for tightly packed layers where translational degrees are constrained.

This move can also be applied to multi-phase simulations. One of the limiting factors of transfer of molecules between two phases, especially if the phases are dense, is solvent cavitation in one phase and solute insertion in the other phase. The same approach as discussed in the previous paragraph can be applied here. Instead of swapping one molecule at a time between phases, a pair of two different molecules in two different phases is selected and their identities are exchanged. Such approach eliminates the need to transfer larger molecules and only swaps of smaller components are performed followed by the identity exchange moves. This technique has been successfully applied in a number of studies including chromatography [30] and oligomer blend phase behavior [31].

Chapter 2

First Principles Monte Carlo Simulations of Reaction Equilibria in Compressed Vapors

2.1 Introduction[1]

Studying the behavior of complex chemical systems undergoing various multiphase reaction equilibria is of crucial importance in both chemistry and chemical engineering. Fundamental understanding of system composition and properties is critical for optimization and design of new processes involving chemical reactions. However, accurate experimental measurements of reaction equilibria can be challenging for extreme conditions (temperature and pressure), hazardous compounds, or nano-confinement. These challenges necessitate development and improvement of predictive modeling techniques for such reactive systems.

Existing computational approaches to study reaction equilibria can be grouped into two general categories: static electronic structure calculations and trajectory-based approaches [33]. Electronic structure calculations with their accurate energetics are usually

[1]This chapter describes the outcome of a research project carried out by E. O. Fetisov in collaboration with I-F. W. Kuo, C. Knight, J. VandeVondele, T. V. Voorhis and advised by J. I. Siepmann. A report on this project has been published [32]. Reprinted with permission from [32]. Copyright 2016 American Chemical Society.

the method of choice, but difficulties arise from the treatment of solvation environments (limited to continuum solvation models or inclusion of a few explicit solvent molecules) and computation of molecular partition functions when multiple conformations make significant contributions. In the group of trajectory-based methods, there are three approaches. Molecular dynamics (MD) methods can be used in conjunction with a first principles description or a reactive force field to represent inter- and intramolecular interactions. The main drawback of MD methods is that reactive processes along with transfer between phases are usually associated with relatively high free energy barriers and, hence, the rates are often beyond accessible time scales. Although reactive force fields [34, 35] allow one to access longer time scales than first principles MD approaches, the former often require explicit parameterization for new systems to be studied [36]. Reactive Monte Carlo (RxMC) approaches [37] can overcome the kinetic limitations through the introduction of special RxMC moves that attempt the direct conversion of molecular species. However, RxMC has two major drawbacks: (i) a set of chemical reactions has to be pre-specified and formation of chemical compounds is limited to those appearing in the reaction set, and (ii) the ratios of molecular partition functions need to be known for all chemical reactions, but may not always be available or applicable in condensed phases. Thermochemical kinetics simulations evolve the system based on rates of chemical reactions that have to be provided as input parameters, and treat the environment via macroscopic equation-of-state approaches.

In this work, we present the reactive first principles Monte Carlo (RxFPMC) method that combines concepts from MD and RxMC approaches. The main idea is simple but conceptually novel in the field of MC simulations. Every molecule in the system is viewed as an aggregate of individual atoms and not as a single entity, and all MC moves are performed on atoms or aggregates of atoms. That is a viewpoint also taken by MD using a first principles description or reactive force fields. In RxFPMC, reactive events are sampled by special MC moves involving the exchange of atoms between aggregates (i.e., molecules) or the transfer of an atom from one aggregate to another. As a first application, we choose a mixture containing dioxygen (O_2), dinitrogen (N_2), and other N_xO_y species. This system is an attractive test case because, in addition to the main reaction of nitric oxide formation ($N_2 + O_2 \rightleftarrows 2\ NO$), further oxidation reactions (e.g. $2\ NO + O_2 \rightleftarrows 2\ NO_2$) can take place, particularly for oxygen-rich mixtures. The

N:O mixture is especially important in atmospheric chemistry where NO formation happens during lightning storms [38], industrial chemistry [39], combustion [40, 41] and rocketry [42] sciences, and astrochemistry [43]. The new RxFPMC framework provides a consistent approach to predict and understand the behavior of chemically reactive systems in highly non-ideal environments.

2.2 Reactive First Principles Monte Carlo Methodology

To provide perspective on the RxFPMC method, we start this section with a brief description of the RxMC method introduced in the early 1990's in three flavors by Shaw [5], Smith and Triska [6], and Johnson *et al.* [7]. The main idea behind RxMC simulations is to introduce a new reactive move in addition to conventional MC moves (e.g, translations and rotations of molecules). During such a reactive move, reactant molecules are artificially converted into products or vice versa according to stoichiometry, i.e., the number of molecules can change through insertion/deletion of the respective molecules. The acceptance rate for such an RxMC move is as follows [37]:

$$P^{\xi} = \min\left\{1, \left(\frac{p^0 V}{k_{\mathrm{B}} T}\right)^{\bar{\nu}\xi} \exp\left(-\frac{\Delta U}{k_{\mathrm{B}} T}\right) K^{\xi} \prod_{i=1}^{s} \left[\frac{(N_i^*)!}{(N_i^* + \nu_i \xi)!}\right]\right\} \tag{2.1}$$

where ξ is the reaction direction ($+1$ and -1 for forward and backward reaction, respectively); p^0, V, and $\bar{\nu}$ are the standard pressure, system volume, and the total change in the number of molecules during the reaction step, respectively; K is the standard-pressure ideal-gas equilibrium constant; ΔU is the change in the potential energy caused by the move; s, N_i^*, and ν_i are the number of molecules participating in the reaction, the number of molecules of type i prior to the reactive move, and the stoichiometric coefficient for this molecule type, respectively. Equation 2.1 requires K as an input parameter and its value is obtained using either known ideal-gas chemical potentials for each molecule type (e.g., from JANAF tables [44]) or using statistical mechanics to calculate ideal-gas molecular partition functions [1]. The reliance on ideal-gas K values for pre-specified reactions leads to the limitations mentioned previously.

The RxFPMC methodology views a reactive system as a collection of independent atoms (or nuclei) that can aggregate to form molecules. To describe the energetics,

we choose Kohn-Sham density functional theory (KS-DFT) due to its robustness and speed [45]. However, switching from a molecular to an atomic description changes the types of degrees of freedom, i.e., rotational and vibrational degrees of freedom for molecules are converted into translational degrees of freedom for atoms. Associated with this switch, the types of MC moves that can be applied to the system change from a case where translational, rotational, and conformational moves are applied on molecules and volume changes involve scaling of center-of-mass coordinates to one where, in the simplest form, translational moves are only applied to atoms and volume changes involve scaling of atomic coordinates. Similarly, the calculation of the pressure switches from molecular virial (only intermolecular forces between molecules) to atomic virial. Thus, a set of test simulations is performed to validate that molecular and atomic representations yield indeed the same answer for prototypical systems where all atoms belong to molecules. Results for this test are described in the next section.

The main challenge for RxFPMC is to find a set of MC moves that can efficiently sample the aforementioned clustering events which, in turn, are chemical reactions. Over the years, various MC algorithms have been developed for strongly aggregating systems, including algorithms that translate or rotate all particles belonging to a cluster together as a group [3] and that sample evaporization/condensation events (i.e., a particle is removed or added to a cluster) [12] and even the transfer of a particle from one cluster to another [13]. While these algorithms were developed to deal with strongly associating molecules, they can be also applied to individual atoms ensuring good statistical sampling. Another useful MC move is an identity exchange of two (or more) different chemical atoms. If the exchanged atoms belong to two different molecules, then the result is a chemical reaction involving two reactant and two product species. If the exchanged atoms belong to the same molecule, then the result is an isomerization. Using configurational-bias Monte Carlo approaches [46] it is also possible to adjust atomic distances (i.e., bond lengths and bend angles) during such an exchange move or to even carry out an identity exchange involving different numbers of atoms (e.g., an H atom could be exchanged with a CH_3 group) [27]. As will be shown in the next section, unbiased N:O atom exchange moves and translations of individual atoms are sufficient for the NO system investigated here.

It is clear that RxFPMC overcomes the limitations of the RxMC method mentioned

above. First, RxFPMC does not require one to specify a set of allowed reactions due to the atomic nature of the system and sampling of atomic aggregates is constrained only by the Boltzmann weights of system configurations. Second, the RxFPMC methodology does not rely on ideal-gas intramolecular partition functions that may lead to systematic errors for highly non-ideal environments where the environment modifies the molecular partition functions (e.g., via the formation of a hydrogen bond or strong adsorption to a substrate). However, it needs to be emphasized that RxFPMC is designed to circumvent the kinetic limitations characteristic to MD simulations and, hence, RxFPMC provides only the equilibrium distribution of reactants and products. In addition, the present implementation samples from the classical system partition function because including nuclear quantum effects via path integral approaches would significantly add to the expense. In contrast, RxMC can utilize quantum molecular partition functions (at least for molecules with limited conformational flexibility).

2.3 Computational Details

The RxFPMC simulations are performed with the CP2K simulation package [47] which solves the KS-DFT equations with the Gaussian plane wave method as implemented in the Quickstep module [48]. Most of the simulations use the BLYP functional [49, 50] with the third-generation Grimme dispersion correction (D3) [51] along with a triple-zeta, double polarization basis set [52], GTH pseudopotentials [53, 54], and a plane wave cut-off at 600 Ry. For one state point, additional simulations with the M06 [55] and rVV10 [56, 57] functionals are carried out to assess the accuracy of the predictions. For the compressed vapor phase, only system sizes and atomic compositions are considered that lead to an even number of electrons and the spin multiplicity of the entire system is restricted to 1, but employing spin-polarized (unrestricted) KS-DFT. Constraining the overall spin multiplicity to 1 is a reasonable assumption for bulk systems, i.e., there are enough molecules in the system to allow the individual molecules to adjust their spin state as needed. Two triplet O_2 molecules that are sufficiently far apart constitute a singlet system, there will be two spin up electrons on one molecule and two spin down electrons on the other. Similarly, two NO molecules can adopt the spin state as needed, i.e., both spin up or spin down when the two NO molecules are formed from the reaction

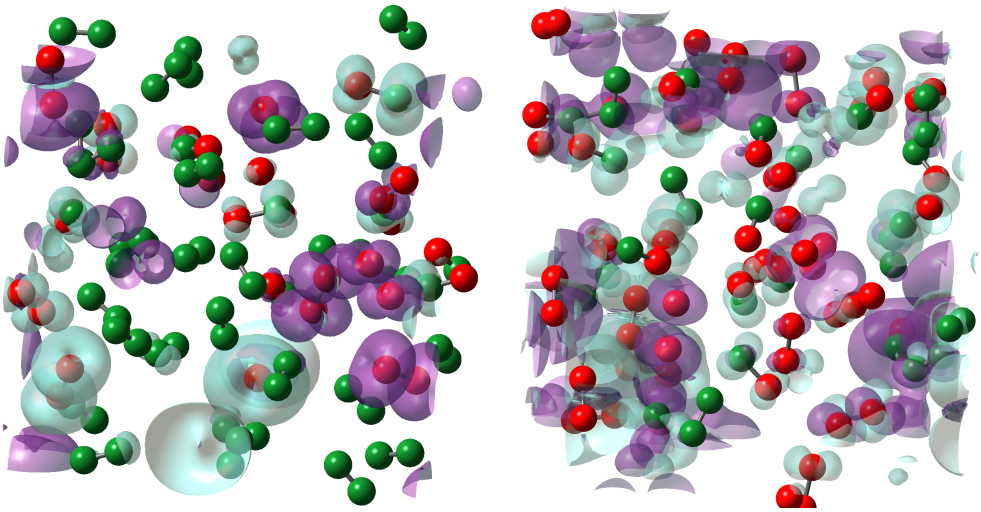

Figure 2.1: Snapshots of 96-atom systems with N:O ratios of 2:1 (left) and 1:2 (right). Nitrogen and oxygen nuclei are shown as green and red spheres, and the purple and cyan isosurfaces represent local spin density (LSD) of different signs.

of a triplet O_2 molecule and a singlet N_2 molecule. Snapshots illustrating the local spin densities for the 2:1 and 1:2 N:O systems are shown in Figure 2.1. All N_2 molecules are found to be singlets without appreciable local spin density, whereas the other species exhibit significant local spin density, e.g., triplet O_2 molecules.

System sizes consisting of 96 and 192 atoms are used to confirm that the RxFPMC approach converges to equilibrium compositions irrespective of the starting compositions (see Figure 2.3). Increasing the system size from 96 to 192 atoms may lessen the constraint on speciation caused by setting the overall spin multiplicity to 1, but increases the average computer time per MC cycle by a factor of about 14. Since the species distributions obtained from the initial simulations for 96- and 192-atoms systems are statistically indistinguishable (see Table 2.2), subsequent simulations exploring different ratios of nitrogen to oxygen atoms (N:O = 5:1, 2:1, 1:1, 1:2, and 1:5) are carried out using the smaller system size in the NpT ensemble ($N_{\text{atom}} = 96$, $p = 30$ GPa, and $T = 3000$ K). Equilibration periods run for at least 100 MC cycles (where one MC cycle consists of N_{atom} randomly selected moves), followed by at least 150 MC cycles of production. The MC move probabilities are adjusted to yield one accepted volume and one accepted N:O atom identity exchange move per cycle, and the remainder is

distributed between O and N translational moves according to the atomic composition. The maximum displacement applied for the volume moves is set to yield an acceptance rate of about 20%; for example, 13 and 20 $Å^3$ are used for the 1:1 N:O systems with 96 and 192 particles, respectively (corresponding to about 0.7 and 0.5% of the total volume, respectively). The N:O atom exchange moves yield an acceptance rate of about 10-15%. The maximum displacements for the translational moves are adjusted to yield acceptance rates of about 30% and are about 0.35 and 0.45 Å for nitrogen and oxygen atoms, respectively. To gather satisfactory statistics, 32 independent simulations are carried out at each composition for the 96-atom system. These are used to estimate the 95% confidence interval for the calculated properties.

2.4 Results and Discussion

In order to demonstrate that MC sampling in an atomic representation yields consistent results with a molecular representation, we investigate three types of related systems containing (system A) N_{molec} diatomic molecules with harmonic bonds, (system B) $N_{\text{atom}} = 2N_{\text{molec}}$ atoms where each atom interacts only with its nearest neighbor through a harmonic potential, and (system C) $N_{\text{atom}} = 2N_{\text{molec}}$ atoms where each atom interacts only with another specific atom in the system (not necessarily its nearest neighbor) through a harmonic potential. In addition, each system is probed either as an ideal gas (only harmonic bond potentials for the diatomic and no intermolecular interactions) or as liquid phase of dinitrogen (with additional Lennard-Jones interactions [58]). For system A, the simulations involve translational moves applied to molecules, rotational moves around the center of mass, conformational moves changing the bond length, and volume moves with scaling of center of mass positions. For systems B and C, only translational moves for individual atoms and volume moves with scaling of atomic positions are used. The number of particles used within the acceptance rule for volume changes is set to N_{molec} for system A and to N_{atom} for systems B and C. As can be seen from the numerical data provided in Table 2.1, these simulations yield molar volumes and internal energies that are statistically indistinguishable for systems A, B, and C and independent of the initial volume used. The small increase in the internal energy of the ideal gas from $RT/2$ is due to rovibrational coupling.

Table 2.1: Equilibrium Molar Volumes and Molar Internal Energies for an Ideal Diatomic Gas at $T = 273.15$ K and $p = 1$ bar and for Liquid Dinitrogen at $T = 73.15$ K and $p = 10$ bar.[a]

	N	$V_{\text{initial}}p/RT$	$\langle\bar{V}\rangle p/RT$			$\langle\bar{U}\rangle/RT$		
			System A	System B	System C	System A	System B	System C
IG	100	0.057	1.000_4	0.998_6	1.000_5	0.51_1	0.50_1	0.51_2
	100	7.16	0.999_5	0.999_5	1.001_6	0.50_1	0.51_2	0.51_1
N_2	1000	0.0007	0.0016_1	0.0017_2	0.0016_3	-2.13_3	-2.15_4	-2.15_4
	1000	0.099	0.0015_3	0.0016_2	0.0016_2	-2.15_4	-2.16_4	-2.13_5

[a] The subscripts denote the 95% confidence intervals. The parameters for the harmonic potential are $k_{\text{f}} = 83$ kJ mol^{-1} Å^{-2} and $r_0 = 1.1$ Å.

Switching to the reactive NO system, we first address whether the RxFPMC method with the selected set of MC moves (atom identity exchanges, atom translations, and volume exchanges) can reach equilibrium distributions of species (independent of initial composition) and whether RxFPMC allows for faster sampling than FPMD. In order to answer these questions, we track the evolution of the molar fraction of NO molecules, x_{NO}, as function of runtime (or number of MC cycles or MD steps). To this extent, we investigate three system compositions for a 192-atom system (N:O ratios of 1:1, 2:1, and 5:1) with initial configurations that contain either only N_2 and O_2 molecules or only NO and N_2 molecules. These initial configurations are taken from a non-reactive Monte Carlo simulation (i.e., using a molecular representation). In the RxFPMC approach (and, of course, also in FPMD), molecules are simply aggregates of atoms and we use a distance cutoff at 1.6 Å to determine whether any two atoms belong to the same molecule. Here we follow the Stillinger approach [59] and use a geometric cut-off criterion to determine whether any two atoms belong to the same aggregate. Such a criterion does not restrict the size of aggregates (i.e., the number of atoms in a molecule). The N–N, O–O, and N–O radial distribution functions (RDFs) for five N:O ratios are shown in Figure **??**. The positions of the first peak are 1.08, 1.23, and 1.12 Å for the N–N, O–O, and N–O RDFs, respectively. The first minimum for the O–O RDFs is found at 1.60 Å. At this location, the N–N and N–O RDFs yield average values for the five compositions that are about 0.0011 and 0.06, respectively. Thus, 1.60 Å is a suitable value that can be applied as the geometric cutoff criterion for N–N, N–O, and O–O

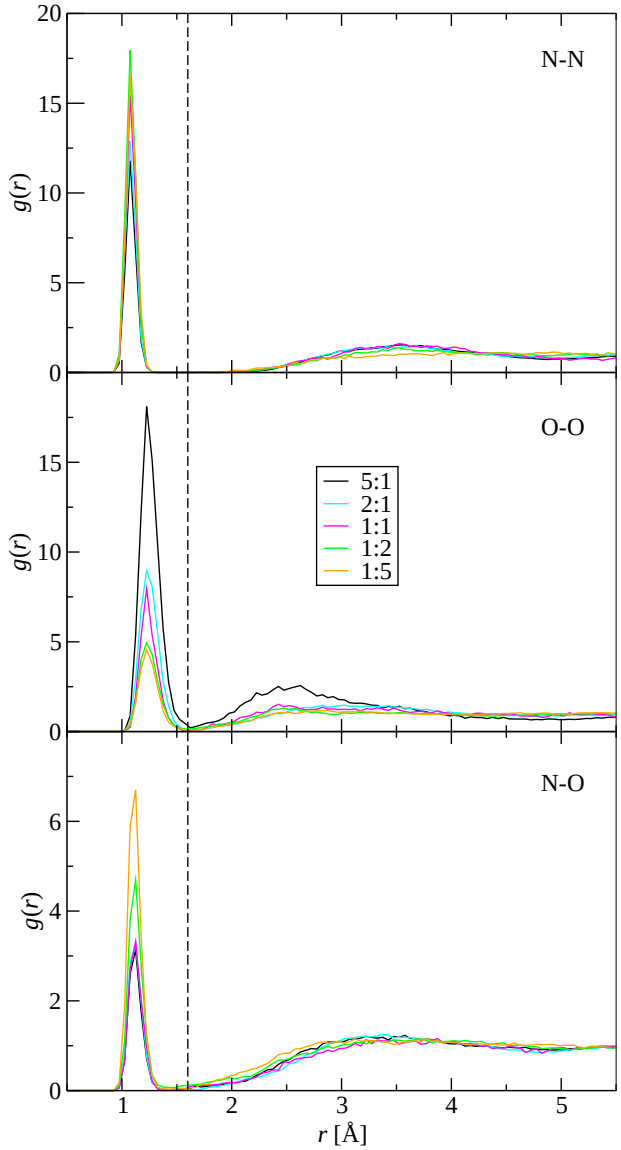

Figure 2.2: Nitrogen–nitrogen (top), oxygen–oxygen (middle), and nitrogen–oxygen (bottom) radial distribution functions obtained for systems with N:O ratios of 5:1 (black), 2:1 (cyan), 1:1 (magenta), 1:2 (green), and 1:5 (orange). The vertical dashed lines show the cutoff criterion for aggregation at $r = 1.60$ Å.

atom pairs.

The evolutions of x_{NO} for the 12 different runs are illustrated in Figure 2.3. It is evident that the RxFPMC simulations rapidly converge to equilibrium x_{NO} values

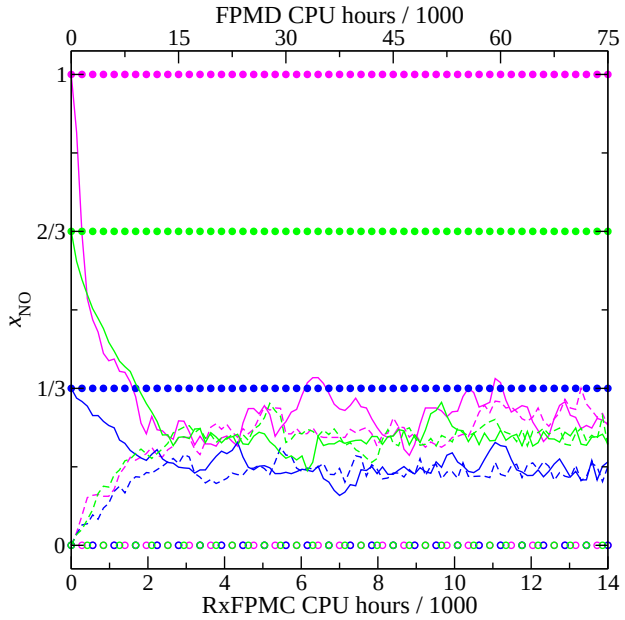

Figure 2.3: Evolution of instantaneous values for the molar fraction of nitric oxide with runtime. The magenta, green, and blue colors denote 192-atom systems with N:O ratios of 1:1, 2:1, and 5:1, respectively. Each RxFPMC simulation consists of 100 MC cycles (taking a total of 14000 CPU hours, see x-axis scale at bottom) and each FPMD simulation covers 60 ps (taking a total of 75000 CPU hours, see x-axis scale at top). The solid lines and filled circles correspond to RxFPMC and FPMD trajectories, respectively, that started with all oxygen atoms being present in the form of NO molecules and the excess nitrogen atoms in the form of N_2 molecules. The dashed lines and open circles symbols correspond to RxFPMC and FPMD trajectories, respectively, that started with all atoms being present in the form of N_2 and O_2 molecules.

(with the highest $\langle x_{NO} \rangle$ for the N:O ratio of 1:1) irrespective of the initial speciation of the system. For all three N:O ratios, equilibrium is reached within about 25 MC cycles (equivalent to a total of 3500 CPU hours running on 192 cores of Intel Haswell E5-2680v3 processors; i.e., each MC move takes on average 14 s of wallclock time). It needs to emphasized that the relatively similar bond lengths for N_2, O_2, and NO (see below) enable a fairly high acceptance rate for the atom identity exchanges (about 10-15%). In contrast, the six 60-ps FPMD trajectories do not yield a single reactive event even at the rather extreme conditions used here ($T = 3000$ K). Due to the high temperature, the FPMD simulations are carried out with a timestep of 0.25 fs (i.e.,

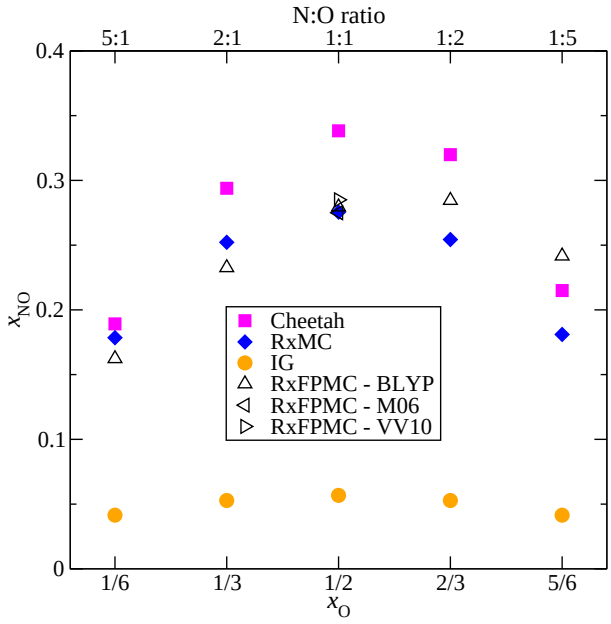

Figure 2.4: Equilibrium molar fractions of nitric oxide versus the oxygen content in the system. Magenta squares, blue diamonds, and orange circles denote predictions using the Cheetah thermochemical code [64], the force-field based RxMC data of Smith and Triska [6], and using ratios of ideal-gas molecular partition functions, respectively. The up, left, and right triangles denote RxFPMC data obtained with the BLYP-D3, M06, and rVV10 functionals. Statistical errors are smaller than the symbol size.

each timestep takes about 6 s of wallclock time). The reason for the wallclock time per MC move being longer than that per MD timestep is that the MC moves can lead to more significant changes of the configuration and, hence, require more SCF steps for convergence. Similar observations regarding the number of SCF steps have been made previously for non-reactive first principles MC simulations [60, 61]. It should also be noted that other recent FPMD studies utilized either a strong electric field [62] or a virtual piston [63] to initiate reactive events.

The NO formation reaction is endothermic by about 180 kJ/mol and, at normal conditions, the equilibrium is shifted very much to the reactant side. In Figure 2.4, the equilibrium x_{NO} values obtained with various approaches are compared (numerical values are provided in Table 2.2). Using ideal-gas molecular partition functions [1] for the $N_2 + O_2 \rightleftarrows 2$ NO reaction yields x_{NO} values ranging from 4.1% to 5.7% for

N:O ratios of 5:1 and 1:1 that, by construction, are symmetric with respect to the oxygen content, x_{O}, and the N:O ratio. In contrast, when intermolecular interactions are taken into account with various approximations (Cheetah using an equation-of-state approach [64, 65], RxMC using a force field, and RxFPMC using KS-DFT), then the equilibrium x_{NO} values are found to be about a factor of five larger than the ideal gas data. This very substantial increase is caused by substantially stronger NO–NO interactions compared to O_2–O_2 and N_2–N_2 interactions [66]. The stronger NO–NO interactions are due to NO's substantial permanent dipole ($\mu = 0.16$ D) [67] and higher polarizability [68], whereas N_2 and O_2 possess only permanent quadrupole moments. On average, the force-field based RxMC simulations of Smith and Triska [6] yield the smallest enhancement of x_{NO}, whereas the Cheetah thermochemical code [64, 65] yields the largest enhancement.

The other important feature is that the RxFPMC simulations and the Cheetah predictions exhibit significantly asymmetric x_{NO} values with a larger enhancement for higher oxygen content or smaller N:O ratio. For example, the x_{NO} values at a 1:5 N:O ratio are larger than those at a 5:1 ratio by factors of 1.5 and 1.14 for RxFPMC and Cheetah, respectively, whereas this factor is only 1.01 for RxMC. For the RxFPMC simulations, the enhancement factors with respect to the ideal-gas values are monotonically increasing and are 3.9, 4.4, 4.9, 5.4, and 5.8 for N:O ratios of 5:1, 2:1, 1:1, 1:2, and 1:5, respectively. This asymmetric enhancement will be discussed further below.

For the 1:1 N:O ratio, RxFPMC simulations are carried out for the BLYP-D3, M06, and rVV10 functionals. Agreement between the three functionals is very good with x_{NO} values of 0.285 ± 0.003 (rVV10), 0.279 ± 0.002 (BLYP-D3), and 0.275 ± 0.003 (M06). This order is also consistent with the reaction energies obtained for isolated molecules that are +179.8, +181.2, and +184.8 kJ/mol, respectively. These values obtained with the pseudopotential approach of CP2K are in good agreement with those obtained from all-electron calculations (see Table 2.3).

One approximation used for the RxFPMC simulations (but not for the RxMC and Cheetah data) is that nuclear quantum effects are neglected in the sampling. To assess the significance of this approximation, the ideal-gas equilibrium constant is calculated using either a quantum-mechanical or a classical-mechanical description of the vibrational partition function. It is found that using the classical description increases x_{NO} by

Table 2.2: Equilibrium Molar Fractions for the Five Different Species Observed in the RxFPMC Simulations.[a]

N:O ratio	Method	NO	O_2	N_2	NO_2	N_2O
5:1	RxMC	0.1785_4	0.0774_4	0.7441_4	–	–
	Cheetah	0.1892	0.0696	0.7363	0.0008	0.0040
	RxFPMC(BLYP)-96	0.162_4	0.085_2	0.746_2	0.001_1	0.006_2
	RxFPMC(BLYP)-192	0.157_10	0.090_5	0.749_5	0.000_1	0.004_3
2:1	RxMC	0.2522_5	0.2072_5	0.5406_5	–	–
	Cheetah	0.2939	0.1816	0.5146	0.0028	0.0069
	RxFPMC(BLYP)-96	0.232_2	0.209_2	0.554_2	0.003_2	0.002_1
	RxFPMC(BLYP)-192	0.229_8	0.216_4	0.551_4	0.002_2	0.002_2
1:1	RxMC	0.2755_4	0.3623_4	0.3623_4	–	–
	Cheetah	0.3383	0.3242	0.3233	0.00615	0.0080
	RxFPMC(BLYP)-96	0.279_2	0.351_2	0.364_2	0.004_2	0.002_1
	RxFPMC(BLYP)-192	0.269_9	0.358_5	0.366_5	0.004_3	0.003_2
	RxFPMC(M06)	0.275_3	0.354_3	0.372_3	0.0043_2	0.003_2
	RxFPMC(rVV10)	0.285_3	0.346_2	0.368_2	0.006_2	0.002_2
1:2	RxMC	0.2543_4	0.5395_4	0.2062_4	–	–
	Cheetah	0.3199	0.4993	0.1650	0.0091	0.0065
	RxFPMC(BLYP)-96	0.285_3	0.517_2	0.184_2	0.010_2	0.004_1
1:5	RxMC	0.1810_4	0.7428_4	0.0762_4	–	–
	Cheetah	0.2149	0.7200	0.0531	0.0090	0.0027
	RxFPMC(BLYP)-96	0.241_5	0.705_3	0.046_2	0.006_1	0.002_1

[a] The subscripts denote the 95% confidence intervals. The RxMC results are taken from [6] and the subscripts denote the precision of the data quoted by these authors. The thermochemical distributions are obtained using the Cheetah 8.0 thermochemical code which employs an equation of state based on exp–6 and dipolar interactions between molecular gas products [64, 65]. The parameters for the equation of state utilized in Cheetah are obtained by fitting to available experimental data as well as simulation results, e.g., for dioxygen about 95% of the data points are sourced from experimental measurements

a factor of less than 1.02 (larger than unity because the zero-point energy is smallest for NO). That is, the neglect of nuclear quantum effects causes a shift that is approximately similar to the statistical error of the RxFPMC simulations. It certainly helps that the NO system does not contain any hydrogen atoms and is studied at high temperature.

The molar fractions of the three diatomic species are compared in Figure 2.5. The data for RxMC [6], Cheetah [64], and RxFPMC simulations are in very good agreement

Table 2.3: Calculated Compound Energies (in Hartree) and Reaction Energy (in kJ/mol) for the $N_2 + O_2 \rightleftarrows 2$ NO Reaction Using Different Functionals and Software Packages.

Functional/Software	N_2	O_2	NO	ΔE
BLYP/CP2K	-19.86320663	-31.88308944	-25.83863239	+181.24
BLYP/Gaussian	-109.5502456	-150.3696483	-129.9255044	+180.86
M06/CP2K	-18.70286776	-30.23584499	-24.43416101	+184.81
M06/Gaussian	-109.4944980	-150.2947843	-129.8599762	+182.03
rVV10/CP2K	-19.89473252	-31.90665637	-25.86644814	+179.83
VV10/QChem	-109.7249145	-150.5760133	-130.1164694	+178.52

with the expected shift in population from N_2 being the major species (with $0.736 \leq x_{N_2} \leq 0.746$) at the 5:1 N:O ratio to O_2 being the major species (with $0.705 \leq x_{O_2} \leq 0.743$) at the 1:5 N:O ratio. As an aside, this shift in population away from N_2 with the shortest bond length (see below) is also reflected in a decrease of the specific density from a value of 1194 ± 1 kg/m^3 for the mixture with the 5:1 N:O ratio to 1078 ± 1 kg/m^3 for the mixture with the 1:5 N:O ratio.

Most importantly, the cluster analysis of the RxFPMC simulations also yields small, but statistically significant molar fractions for the NO_2 and N_2O triatomic molecules (see Figure 2.5 and Table 2.2). The RxFPMC simulations do not show dissociation into atoms nor formation of homoatomic O_3 and N_3 species or species with four or more atoms. The presence of triatomic species demonstrates that the RxFPMC simulations with the current move set are indeed capable of generating new molecules, whereas the RxMC simulations are limited to only those species appearing in the reaction set (i.e., only the diatomic molecules for the NO system). Since we do not detect significant fractions of atomic species, the formation/destruction of the triatomic molecules involves mostly groups of six atoms that rearrange from three diatomics to two triatomics.

The Cheetah simulations also yield NO_2 and N_2O as the only non-diatomic species that are present in molar fractions greater than 2×10^{-4}, but much better statistics for the thermochemical code allow for detection of O_3, O, and N species. With the exception of the oxygen-rich system (N:O ratio of 1:5), where the x_{O_3} exceeds 10^{-4}, all of these minor species are found individually only in molar fractions smaller than 10^{-4} in the Cheetah simulations. Averaging over the five compositions, the RxFPMC

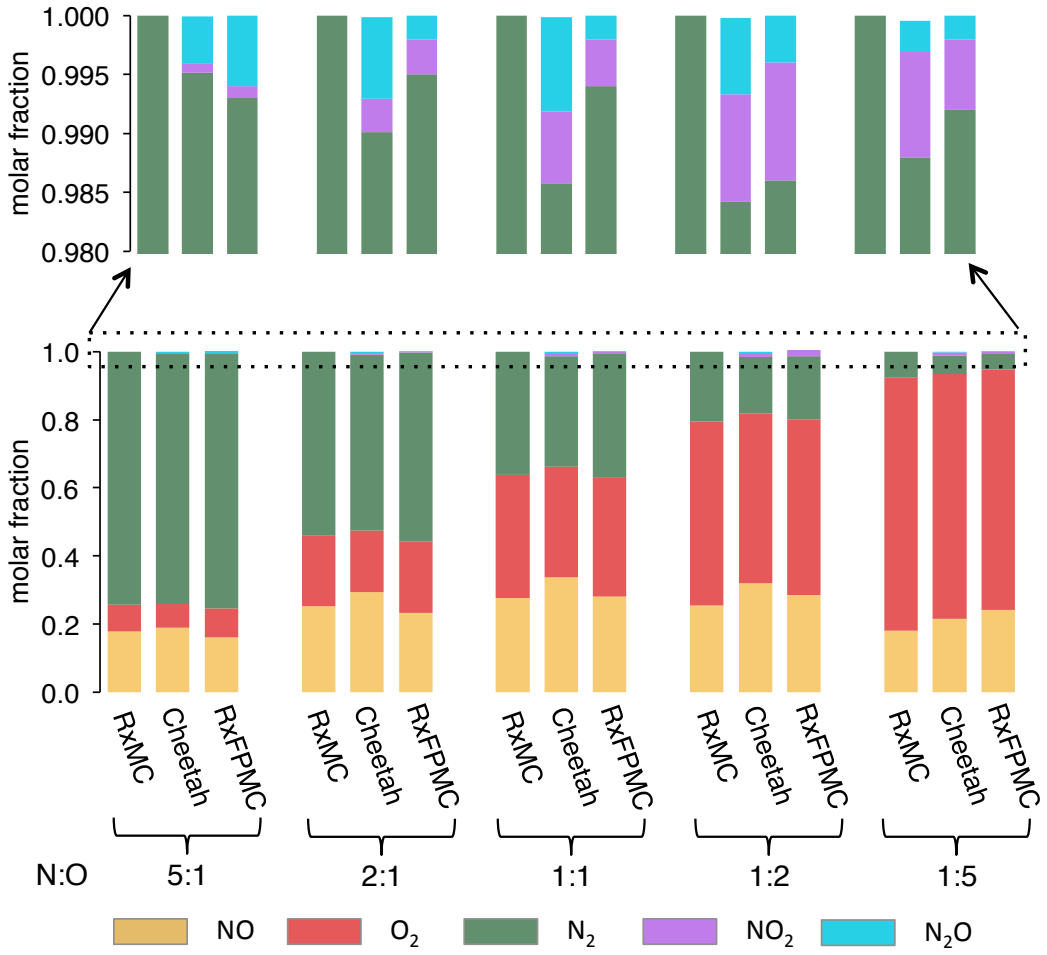

Figure 2.5: Equilibrium molar fractions for NO, O_2, N_2, NO_2, and N_2O obtained from RxMC [6], Cheetah [64], and RxFPMC (BLYP-D3) simulations. The upper part provides a magnified view of the molar fractions of the two triatomic species.

simulations yield a total molar fraction for the two major triatomic species of 0.008 compared to a value of 0.011 for the Cheetah simulations. The RxFPMC and Cheetah simulations also yield about five times more N_2O than NO_2 for the nitrogen-rich system with the N:O ratio of 5:1, and about three times more NO_2 than N_2O for the oxygen-rich system with the N:O ratio of 1:5. With regards to the two isomers possible for the triatomic species, the RxFPMC simulations yield only the N–N–O isomer for N_2O, and the N–O–O isomer constitutes less than 5% of the NO_2 molecules, i.e., almost all are

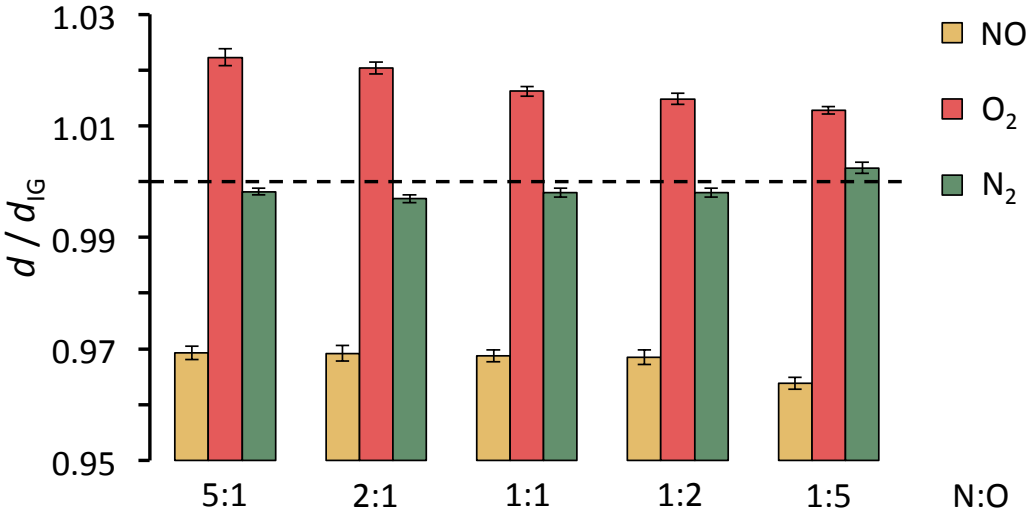

Figure 2.6: Ratios between average bond lengths of the diatomic molecules in the bulk phase and for an isolated molecule at $T = 3000$ K for the five compositions studied with the BLYP-D3 functional.

present as the O–N–O isomer. The data for the BLYP-D3, M06, and rVV10 functionals are in agreement with respect to the fraction of triatomic molecules and their isomers for the 1:1 N:O mixture (see Table 2.2).

Overall, the agreement with the Cheetah data is very encouraging. The present simulations of reactive equilibria are carried out at high temperature and pressure where repulsive, first-order electrostatic, and induction interactions between molecular species are more important than dispersive interactions. Furthermore, the focus is on equilibrium distributions and not on reaction rates where sampling of transition states would be important. These factors may mask deficiencies of the KS-DFT and also explain the good agreement for data obtained with different functionals, whereas vapor–liquid equilibria are known to be very sensitive to details of the electronic structure calculations [69].

To provide an explanation of the asymmetric molar fraction for NO observed in the RxFPMC simulations (see Figure 2.4), we turn our attention to microscopic-level information that can be obtained only from the RxFPMC simulations but not from an equation-of-state based approach. Data for the average bond lengths of the diatomic

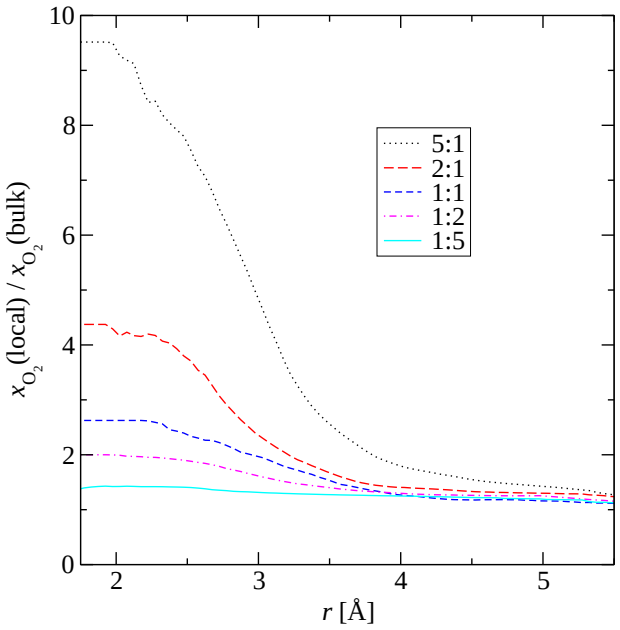

Figure 2.7: Local enhancement for O_2 molecules surrounding NO molecules calculated for the systems with the five N:O ratios studied using RxFPMC and the BLYP-D3 functional.

molecules are presented in Figure 2.6. Compared to isolated molecules, the environment provided by the highly compressed gas phase leads to a shortening of the NO bond length by about 3%, whereas the O_2 bond is lengthened by about 2%. As expected, the effect on the N_2 triple bond is very small. Taking bond length as a proxy for bond strength, the compressed-vapor environment leads to a stabilization of the NO radical, whereas O_2 is destabilized. This finding supports higher x_{NO} values for the oxygen-rich mixtures (N:O ratios of 1:2 and 1:5) than for the corresponding nitrogen-rich mixtures. It is also consistent with the observation that x_{O_2} for the 1:5 N:O mixture is somewhat smaller (by a factor of 1.06) than x_{N_2} for the 5:1 N:O mixture. It should also be noted that the extent of the lengthening of the O_2 bond is found to decrease slightly with increasing oxygen content.

From an analysis of the various radial distribution functions and the corresponding number integrals another structural feature emerges that supports the higher molar fraction of NO in oxygen-rich environments. Combining data from various number integrals, the local molar fraction enhancement can be computed that provides information

on preferential solvation. For the NO system, we find such an enhancement for the solvation of NO molecules by O_2 molecules (see Figure 2.7). That is, the local surrounding of an NO molecule is occupied on average by significantly more O_2 molecules than one would expect from the bulk composition and random mixing. As for other systems [70], this relative enhancement is largest for the O_2-poor mixture, i.e., the 5:1 N:O mixture with $x_{O_2} = 0.085$. The enrichment is greatly diminished for the 1:5 N:O mixture with $x_{O_2} = 0.705$ because sufficient O_2 molecules are present to provide NO's preferred solvation environment. This local microheterogeneity also supports the higher x_{NO} values for the oxygen-rich mixtures when NO can be stabilized by solvation with O_2.

In conclusion, a new first principles Monte Carlo approach for simulating reactive equilibria is introduced that does not rely on any set of pre-specified reactions nor on ideal-gas equilibrium constants that hamper conventional RxMC simulations. Using a first principles description of the system's energetics allows one to view molecules as being formed by the strong aggregation of atoms. Special Monte Carlo moves can then be used to efficiently sample the distribution of atoms over molecular species, thereby overcoming the kinetic limitations of first principles MD simulations. Application of the RxFPMC method to nitrogen/oxygen mixtures in the highly compressed vapor phase region shows significant enhancement of NO molecules. Structural analysis demonstrates stabilization of NO and destabilization of O_2 molecules and preferential solvation of NO by O_2. These structural features support an asymmetry in the species distributions and a maximum in the NO concentration when oxygen is present in slight excess. This study illustrates the promise of using RxFPMC to investigate reaction equilibria in highly non-ideal environments.

Chapter 3

Understanding the Reactive Adsorption of H_2S and CO_2 in Sodium-Exchanged Zeolites

3.1 Introduction[1]

Sour natural gas is a complex mixture of methane, carbon dioxide, hydrogen sulfide, water, nitrogen, ethane, and other light hydrocarbons. Gas sweetening refers to the removal of acid gases (H_2S and CO_2) for reducing environmental concerns and corrosion of gas pipelines and other equipment. Traditionally, amine-based alkaline solutions have been used to capture these acidic impurities [72, 73]. However, this separation process is very energy-intensive and generally requires a large footprint. Recently, ionic liquids [74, 75] and deep eutectic solvents [76] have been suggested as more sustainable physical solvents for H_2S/CO_2 capture; however, the cost of such absorption-based separations still remain relatively high. Adsorption using molecular sieves, such as cation-exchanged zeolites, has been a promising alternative for gas sweetening in general and selective H_2S (over CO_2) removal in particular [77, 78, 79, 80].

[1] This chapter describes the outcome of a research project carried out by E. O. Fetisov in collaboration with M. S. Shah, C. Knight, M. Tsapatsis and advised by J. I. Siepmann. A report on this project has been published [71]. Reprinted with permission from [71]. Copyright 2018 Wiley-VCH.

H_2S and CO_2 can react chemically to form carbonyl sulfide (COS) and H_2O:

$$H_2S + CO_2 \rightleftharpoons COS + H_2O \tag{3.1}$$

The forward reaction is endothermic and the gas-phase equilibrium constant for this reaction,

$$K_c = \frac{[COS][H_2O]}{[H_2S][CO_2]}, \tag{3.2}$$

ranges from a mere $\approx 10^{-6}$ at 20 °C to $\approx 10^{-2}$ at 400 °C [81]. Thus, in the absence of any other agent or environment, conversion to COS can essentially be arrested. However, it is known that adsorption in cation-exchanged zeolites can lead to significant conversion [82], and also absorption in ionic liquids and deep eutectic solvents may potentially shift the equilibrium. COS possesses similar toxicity as H_2S, but its detection and removal is arguably more challenging than that of H_2S and it may leave the natural gas processing facility as "hidden" sulfur. Therefore, it is desirable to suppress the formation of COS from H_2S and CO_2 [83].

In addition to sweetening, dehydration constitutes an important step in natural gas processing. This is done for three main reasons. Firstly, condensation of water vapor into liquid slugs and/or hydrates, can damage pipelines and downstream compression equipment. Secondly, the acid gases are soluble in water and can form an acidic solution that can potentially corrode pipelines. And thirdly, since water has a higher polarity compared to H_2S, use of polar adsorbents for H_2S requires a pretreatment step to significantly reduce the moisture content of the gas stream. However, this dry stream may further exacerbate the problem by shifting the equilibrium in Eq. 3.1 to yield higher COS conversion.

Fellmuth et al. [82] first investigated the effect of confinement on the conversion of acid gases to COS using sodium-forms of the faujasite (FAU) zeolite with low Si/Al ratios (1.1 to 2.7). They observed that the reaction equilibrium shifts toward COS at lower Si/Al ratios which is accompanied by a higher density of charge-compensating Na^+ cations. This behavior was attributed to the dissociative chemisorption of H_2S in these zeolites where SH^- anions can interact strongly with the weakly coordinated cations and protons with the basic oxygen sites of the zeolite framework [82]. In the 1960s, Union Carbide developed metal aluminosilicate molecular sieves, with nominal pore

diameters of 4–5 Å, for the dehydration of sour gases [84]. Subsequently, in the 1990s, Bülow and coworkers proposed new modifications, such as cation-exchange in LTA and FAU zeolites with larger cages ($\approx$ 11 Å) and choice of operating conditions (to benefit from slower kinetics), to suppress COS formation [81, 83, 85]. Since it is clear that confinement can influence the reaction extent, the opportunity exists to select different zeolite framework structures and to tune Si/Al ratio, Al siting, and cation chemistry to achieve sour gas sweetening with minimal conversion to COS. To reach this goal, it is desirable to understand the governing reactive phenomena at atomic length scales in order to provide guidelines for aiding materials design. In this work, we use first principles simulations to study reactive mixtures of H_2S and CO_2 in the zeolite beta (BEA) with medium-sized 7 Å pores [86].

3.2 Computational Details

Recently, we introduced the reactive first-principles Monte Carlo (RxFPMC) method that can be applied to model reactive systems in highly non-ideal environments [32]. RxFPMC is able to overcome both the kinetic limitations hampering molecular dynamics simulations of rare reactive events and the reliance on pre-specified reaction networks and ideal-gas equilibrium constants inherent in conventional reactive Monte Carlo (RxMC) simulations [37]. The main idea behind RxMC is to introduce special MC moves that directly convert molecular species according to stoichiometry of pre-specified reactions where the acceptance probability also includes a term reflecting the ideal-gas equilibrium constant [6, 7]. However, molecular structure and vibrational modes may be significantly altered in systems under confinement, especially, those with strong guest–host interactions. Such changes may limit the applicability of RxMC relying on ideal-gas equilibrium constants. It should be noted that both RxMC and RxFPMC approaches are not limited to single-phase systems (using the canonical and isobaric-isothermal ensembles), but can also be used for multi-phase reactive equilibria with molecule transfer moves in the Gibbs ensemble [61, 87, 88]. Here, we apply RxFPMC to study reactive mixtures with three different H_2S/CO_2 ratios in all-silica and sodium-exchanged zeolite BEA. It should be be noted that RxFPMC (and also RxMC) simulations provide only equilibrium properties and do not provide any insights about the actual reaction

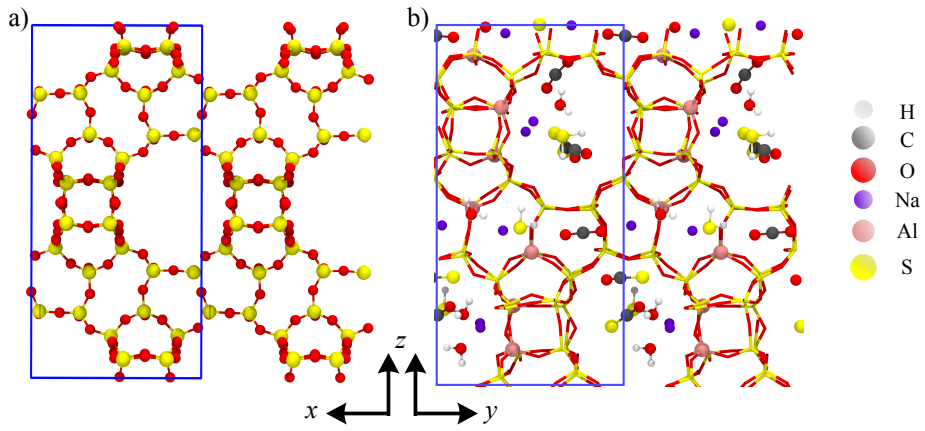

Figure 3.1: a) Two unit cells of all-silica zeolite BEA. b) Equilibrated configuration for the 8:8 H_2S:CO_2 mixture in BEA framework L-3 (Si/Al ratio of 7) with 3 Na^+ pairs.

mechanism (e.g., transition states).

3.2.1 Preparation of Initial Structures

For zeolite BEA (see Figure 3.1), we explore both the all-silica form (Si/Al = ∞) and three different representatives with an intermediate Si/Al ratio of 7. Despite continuing advances in experimental characterization techniques [89], determining the precise location of acidic sites in zeolites remains a challenge. It is usually assumed that the Al distribution follows Löwenstein's rule [90], which empirically states that Al–O–Al triplets are not present in any aluminosilicate framework, but recent calculations suggest possible violations of this rule citeFletcher2017.

The initial structure of the all-silica form of zeolite BEA is taken from the IZA-SC database [91]. In order to generate three structural variants of the Na-BEA zeolite, 8 random Si atoms are replaced with Al atoms following Löwenstein's rule [90]. Then, sodium atoms are placed in the micropore near the Al atoms and, thereafter, the framework structure with the Na^+ cations is allowed to relax with the interactions described at the same level of electronic structure theory as used for the simulations (see below) but maintaining the unit cell dimensions of the all-silica BEA structure. The main difference between the three Na-BEA frameworks is the number of Al pairs occupying

second-nearest neighbor tetrahedral framework sites and correspondingly also the number of Na^+ pairs (with $r_{Na-Na} \leq 6$ Å); these are 3, 2, and 0 for frameworks L-3, L-2, and L-0, respectively. The reason for preserving the unit cell dimension for the current RxFPMC simulations is that allowing for fluctuations in the volume and unit cell parameters would significantly add to the computational expense [92]. To assess that substitution of Si with Al and addition of the Na^+ cations does not lead to significant distortion of the unit cell, optimizations of these three structures are carried out and resulted in only small changes (less than 2%) in the unit cell vectors.

As the last step in the set-up, 16 reactant molecules with H_2S:CO_2 ratios of either 8:8, 4:12, or 12:4 are placed into the BEA frameworks using interactions described by the TraPPE force field [93, 94, 95].

3.2.2 Simulation Details

All simulations are performed using the Monte Carlo modules [32, 60, 61] available in the CP2K software suite [96] .The potential energy of a configuration (defined by the positions of the nuclei) is calculated using Kohn–Sham density functional theory (KS-DFT) with the PBE density functional [97] along with the D3 dispersion correction of Grimme [98], the double-ζ basis set optimized for molecular systems [52], the GTH pseudopotentials [53, 54], and a plane wave cutoff of 400 Ry. The PBE functional has been shown to provide a good energetic description of the flexibility of zeolitic frameworks [99]. To assess the PBE functional (and the other electronic structure parameters) for the H_2S–CO_2 reaction, we compute the ideal-gas equilibrium constant for the H_2S + $CO_2 \rightleftharpoons H_2O$ + COS reaction from the quantum molecular partition functions [1] using the CP2K software with the same electronic structure parameters as for the RxFPMC simulations and from all-electron calculations at the PBE–D3/6-311+G(d,p) level of theory using the Gaussian 09 software [100]. The predicted equilibrium constants of 2.0×10^{-7} (CP2K) and 1.0×10^{-6} (all-electron) at $T = 293$ K are in good agreement with the experimental value of 1.4×10^{-6} [81].

The simulated systems consist of one unit cell of all-silica BEA (with 64 SiO_2 formula units, 192 atoms) or of Na-exchanged BEA with Si/Al ratio of 7:1 (56 Si, 8 Al, 128 O, and 8 Na atoms) and 48 guest atoms with the following three initial reactant ratios: 4:12, 8:8, and 12:4 H_2S:CO_2 molecules. The linear dimensions of BEA's tetragonal unit cell are

$12.632 \times 12.632 \times 26.186$ Å^3. To reduce the computational cost, the current simulations sample the classical partition function of the system (i.e., nuclear quantum effects are not considered). The sampling of the Brillouin zone for the RxFPMC simulations is done only at the Γ-point. Sampling with a $4 \times 4 \times 4$ Monkhorst–Pack k-point mesh is found to yield very similar energy differences between nuclear configurations of the system (within 1 kJ/mol) but at a more than 20-fold increase in computational cost, suggesting that the Γ-point sampling provides good accuracy at a greatly reduced cost.

The RxFPMC simulations are carried out in the canonical NVT ensemble with periodic boundary conditions. The RxFPMC simulations for all three reactant ratios are run at two different temperatures (300 and 450 K), and simulations for the 8:8 ratio are also carried out for three different sitings of the Al atoms at 450 K.

Several different Monte Carlo moves are employed in order to sample the classical-mechanical phase space. First, atom translations are performed on all non-framework atoms (i.e., it is tracked whether an O atom belongs to a sorbate molecule or the BEA framework) and sodium atoms. This atom translation move consists of the random selection of an atom type, the random selection of an atom belonging to this type, and an attempted displacement of this atom. The relative probabilities for these atom translations are chosen equal to the ratio of the numbers of atoms of the specific types. For example, for the 8:8 H_2S/CO_2 mixture in Na-BEA, translations for H, O, C, S, and Na atom types are attempted with the corresponding 2:2:1:1:1 ratios. It should be noted that these atom translations can, in principle, also result in a reactive event, e.g., a proton transfer from H_2S to an acidic oxygen atom of the framework. Secondly, (molecular) translations and rotations on clusters of atoms are performed in order to improve sampling of molecular motion. Clusters are defined as a group of atoms using a simple Stillinger definition [59] with a cutoff radius of 1.8 Å, and the acceptance probability for this move is taken from the work by Orkoulas and Panagiotopoulos [101]. As shown in Figure 3.2, this cutoff value encompasses all chemical bonds (O–H, O–C, S–H, and S–C) and, only rarely, are two water molecules selected as a cluster and displaced together. Maximum displacements for atom translations and cluster translations and rotations are adjusted to yield acceptance rates of about 40%. In order to efficiently sample framework flexibility, a hybrid Monte Carlo (HMC) move is used [102]. During each HMC move, a short microcanonical molecular dynamics (MD)

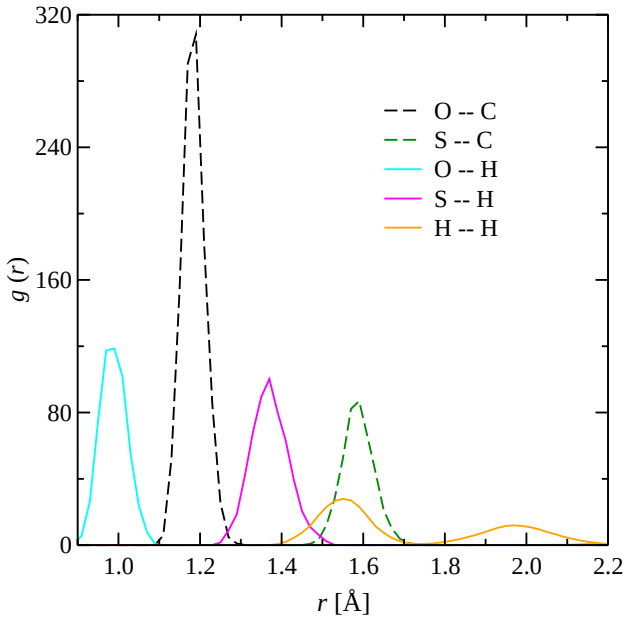

Figure 3.2: Interatomic pair distribution functions relevant for cluster and reactive MC moves computed for the 8:8 H_2S:CO_2 mixture at $T = 450$ K in Na-BEA framework L-3.

trajectory is launched (we use 10 time steps, each time step is 0.25 fs), and the MD sequence is accepted using the conventional MC acceptance rule. Finally, we utilize the special reactive moves discussed in detail in Section 3.2.3. The overall distribution of MC moves is as follows: 15% of atom translations, 20% of cluster translations, 20% of cluster rotations, 20% of HMC moves, and 25% of reactive moves.

For each of the 8 Na-BEA systems (differing in H_2S:CO_2 ratio, temperature, and/or arrangement of Al atoms), 32 independent simulations are carried out. Each simulation consists of at least 10,000 MC moves for equilibration and at least 20,000 MC moves for production. The standard deviations in the computed properties are determined from the 32 independent simulations.

3.2.3 Reactive Moves

The atomic representation of a reactive system in an RxFPMC simulation allows for the use of simple moves (e.g., atom identity exchange) to sample reactive events [32]. While such atom identity exchanges work very well for reactions, where reactants and

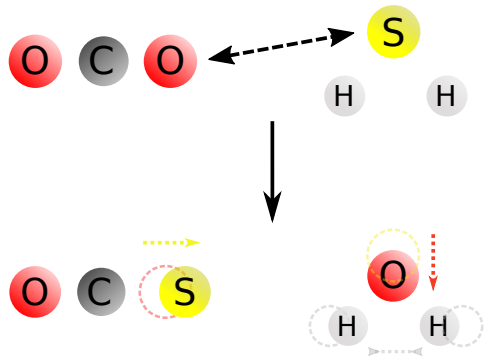

Figure 3.3: Schematic representation of the MC move sampling the reaction between CO_2 and H_2S.

products have similar bond lengths (e.g., $N_2 + O_2 \rightleftharpoons 2\ NO$) or where the distributions of bond lengths are broad at high temperature (as for the N_2/O_2 system at 3000 K [32]), it leads to poor acceptance rates at moderate temperatures ($T < 500$ K) when the chemical bonds belonging to reactant and product molecules are significantly different in length ($\delta r > 0.2$ Å). For the system of interest in this work, the difference between the C–S bond in COS (1.56 Å) and the C–O bond in CO_2 (1.16 Å) is well above 0.2 Å (optimized values for isolated molecules are quoted here, see Figure 3.2 for the actual distributions of bond lengths inside the zeolite). In addition, equilibrium geometries of H_2S and H_2O are considerably different too when it comes to the H–H (1.92 Å and 1.52 Å, respectively) and S(O)–H distances (1.34 Å and 0.96 Å, respectively). A simple exchange of the positions of an O from a CO_2 molecule and an S atom from an H_2S molecule would lead to COS and H_2O product molecules with very unfavorable internal geometries, i.e., a far too short C=S bond length and far too long O-H bond lengths; the acceptance probability for such move would be negligibly small at 300 K.

Therefore, in order to accelerate sampling of the chemical reactions relevant for the acid gas mixtures, we employ a "smart" atom identity exchange move. Every time an exchange of S and O atoms is attempted, new positions for the atoms are picked on the basis of positions of two other atoms in the reactant molecules (see Figure 3.3). Specifically, a new position for the S atom in a product COS molecule is chosen by picking a point along the C–O bond vector (involving the removed O atom) that is

exactly 0.35 Å farther from the C atom than the removed O atom. A new position for the O atom in a product H_2O molecule is chosen by picking a point along the vector from the removed S atom to the midpoint of the two H atoms that is exactly 0.35 Å closer to this midpoint. In addition, new positions for the two H atoms are selected by displacing both H atoms inward along the H-H vector by exactly 0.12 Å. For the reverse reaction, the same rules apply but with the opposite signs for the displacement distances. It should be noted that this procedure yields exactly one internal conformation for the two molecules resulting from the attempted move and, hence, a Jacobian factor is not involved in the acceptance criterion. The specific displacement distances noted above are selected based on short trial runs with different sets of displacements (i.e., they balance the need to select internal distances near the maxima of the relevant distributions with the need to limit changes in the molecular volumes and avoid overlaps with neighboring atoms), and they yield an acceptance probability for the reactive moves of about 4% at $T = 300$ K.

It has to be noted that in order to perform the transformation of internal coordinates during this reactive move, one needs to know to which types of molecules the O and S atoms being exchanged belong. For example, for the forward reaction, we do not pick a random O atom in the system but we pick an O atom connected to a C atom. Similarly, we pick a S atom connected to a H atom, and not to a C atom. This selection rule introduces a bias in the acceptance criterion similar to the combinatorial term in the acceptance criterion for the standard RxMC method [37].This special move does not limit the conversion of CO_2 to COS, but also allows for the subsequent step (H_2S + COS $\rightleftharpoons$ CS_2 + H_2O), because an O atom is picked on the basis of its connectivity to a C atom, and not on the basis of belonging to a CO_2 molecule; that is, a COS molecule can react further with a H_2S molecule to form CS_2.

3.3 Results and Discussion

Short exploratory simulations (2,000 MC moves) at both 300 and 450 K for the 8:8 H_2S:CO_2 mixture in the all-silica form of BEA do not yield any conversion. Although much longer simulations would undoubtedly encounter some reactive events, these exploratory simulations suggest that confinement in the BEA micropores is not sufficient

Table 3.1: Average Conversion of Limiting Reactant, Molar Fraction of Water, Equilibrium Constant, and Internal Energy of Reaction for Different Initial Compositions and Frameworks with Different Al Locations.[a]

T [K]	H_2S:CO_2	L-Al	ξ_L [%]	x_W [%]	K_c	ΔU [kJ/mol]
300	4:12	L-3	26_6	7_3	0.03_1	9_2
300	12:4	L-3	30_8	8_3	0.05_1	8_2
300	8:8	L-3	44_4	22_2	0.62_8	7_1
450	4:12	L-3	35_5	9_3	0.07_3	5_2
450	12:4	L-3	39_3	10_2	0.10_4	5_2
450	8:8	L-3	54_3	27_4	1.38_6	4_1
450	8:8	L-2	34_8	17_5	0.04_2	8_2
450	8:8	L-0	25_6	13_5	0.02_1	11_2

[a] Numbers in subscripts represent standard deviation in the last digit.

to drive the reaction forward and a stronger interaction between the products and the zeolite is required. The lack of appreciable conversion is consistent with experimental findings for other all-silica zeolites [81]. Previously, the effect of confinement on model chemical reactions was studied with standard RxMC simulations. Enhancements up to 40-fold were found for the NO dimerization reaction in porous carbon materials [103, 104], where the increased number density in the pore (compared to the bulk phase) leads to higher conversion for this dimerization reaction with a decrease in the number of molecules in accord with Le Chatelier's principle [105]. However, only small enhancements were observed for reactions with balanced numbers of reactants and product molecules, such as esterification [106], propene metathesis [107], and xylene isomerization reactions [108].

Data for the conversion of the limiting reactant, the molar fraction of water, the equilibrium constant, and the internal energy of the reaction (equal to the reaction enthalpy under the assumption that $\Delta V = 0$) are given in Table 3.1. In Na-BEA, K_c is increased by four orders of magnitude relative to the corresponding gas-phase value. The highest conversion is observed for the equimolar mixture since statistically products are maximized when concentrations of reactants are the same. The conversions for the cases with a 3:1 excess of one of the reactants are lower by a factor of 1.5, and the degree

of conversion is slightly higher (but within the combined statistical uncertainties) for the H_2S-rich mixture as compared to the CO_2-rich mixture. Increasing the temperature from 300 to 450 K yields only a two-fold increase of K_c in Na-BEA compared to a factor of 110 for the ideal-gas phase. It should also be noted that we do not observe formation of any CS_2 molecules. This is somewhat expected since the gas-phase conversion of COS into CS_2 is thermodynamically more prohibitive than of CO_2 into COS and, in addition, COS molecules with their small dipole moment profit more strongly from interactions with the cationic zeolite than the quadrupolar CS_2 molecules.

We also investigate whether the spatial distribution of Al^{3+} atoms in BEA (and consequently the Na^+ arrangement and, specifically, the fraction of Na^+ pairs) impacts the reaction equilibrium (see Computational Details for description of the three framework structures used here). At $T = 450$ K, the Na^+ cations are not very mobile and predominantly occupy positions close to the Al sites. For the 8:8 mixture, we find that the conversion of H_2S decreases from 54 to 34 to 25% with the decrease in the number of Na^+ pairs. This suggests that a larger local concentration of Na^+ cations helps in better product stabilization, and an increase of the number of Na–Na pairs may contribute to the experimentally observed trend of higher conversion for lower Si/Al ratios [81].

Calculation of the internal reaction energies (see Table 3.1) provides the thermodynamic reason for the large increase of K_c and the relative temperature insensitivity found in Na-BEA compared to the small effect of confinement in all-silica BEA. The reaction is strongly endothermic in the gas phase with $\Delta U = 29$ kJ/mol, but interactions of the products (see below) with Na-BEA reduce ΔU by more than 20 kJ/mol. The small ΔU values of ≈ 8 and 5 kJ/mol at $T = 300$ and 450 K, respectively, in Na-BEA lead to strongly increased conversion and a small temperature dependence. In contrast, confinement in all-silica BEA reduces ΔU by less than 2 kJ/mol. Furthermore, ΔU is found to increase for the framework structures with fewer Na^+ pairs.

Structural analysis of the RxFPMC trajectories can provide insight on whether strong interactions of reactants and products with the Brønsted basic oxygen atoms of the framework (and proton transfer) or with the Na^+ cations are primarily responsible for the increased conversion. Analysis of preferential siting around the acidic oxygen (i.e., the ratio of H_2O over H_2S bound to O_{ac} versus the global ratio of these molecules) yields only an enhancement of 1.05. Similarly, the fractions of HO^- and HS^- species are

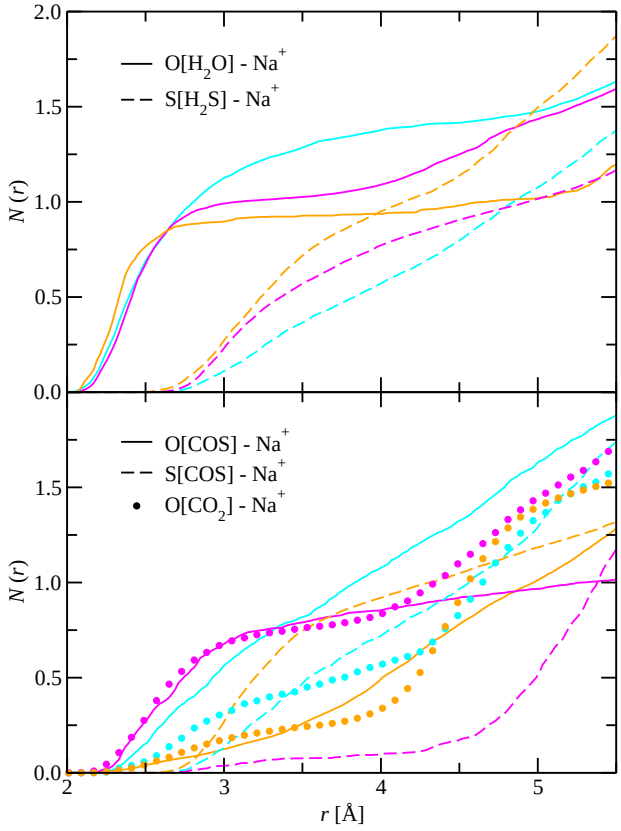

Figure 3.4: Number integrals for Na^+ cations surrounding the O and S atoms of the reactant and product molecules computed for the 8:8 H_2S:CO_2 mixture at $T = 450$ K: (top) coordination for H_2O and H_2S, (bottom) coordination for CO_2 and COS. Cyan, magenta, and orange colors indicate Na-BEA frameworks L-3, L-2, and L-0, respectively, with different numbers of paired Na^+ cations.

found to be vanishingly small. Thus, acid-base chemistry does not play a significant role in the increased equilibrium conversion in Na-BEA but, of course, is likely important for increasing the reaction rate [82].

Analysis of pair distribution functions (PDFs) allows to quantify the coordination environments of the reactant and product molecules. Number integrals (NIs) for Na^+ cations surrounding the O and S atoms of the four different molecules obtained for the 8:8 mixture are shown in Figure 3.4. The Na^+ cations preferentially interact with the H_2O molecules (the peak in the corresponding PDF is at 2.4 Å) and coordination numbers (estimated from the inflection points of the plateau in the NIs) are 1.4, 1.0,

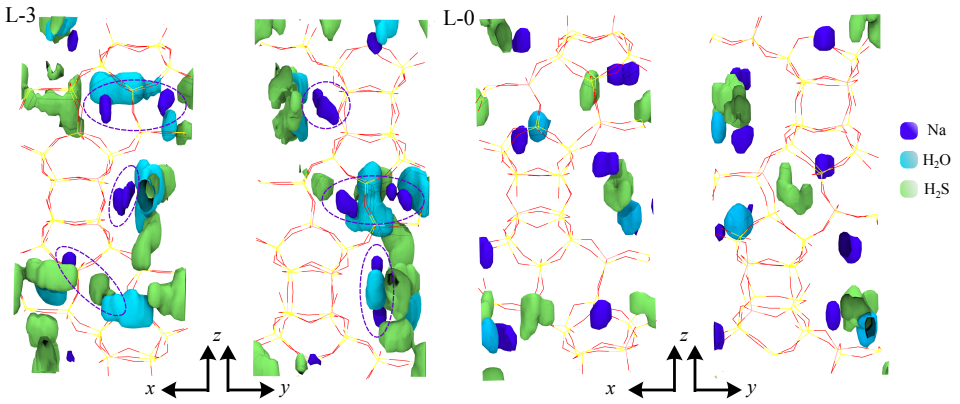

Figure 3.5: Spatial distributions of Na^+ cations, H_2O and H_2S molecules in Na-BEA frameworks L-3 (left) and L-0 (right) with 3 and 0 Na^+ pairs, respectively, per unit cell. The isosurfaces indicate regions with densities larger than twice the corresponding average densities in the micropore. The dashed ellipses highlight the locations of the Na^+ pairs.

and 0.9 for frameworks L-3, L-2, and L-0, respectively. As surmised above, the presence of Na^+ pairs yields enhanced conversion because H_2O molecules can be coordinated by two cations which, in turn, gives a greater degree of product stabilization. In contrast, the NIs for H_2S are shifted to larger distances (peak in the PDF at 3.1 Å), and the coordination number in the different frameworks follow the opposite as for H_2O, i.e., the smallest coordination number of about 0.7 is observed for the framework with the largest number of Na^+ pairs. The strongly enhanced coordination of H_2O by Na^+ cation is due to water's higher dipole moment and oxygen's smaller van der Waals diameter. Previous experimental data and computations also demonstrated preferential adsorption of H_2O over H_2S on Na-exchanged faujasites [78].

Figure 3.5 illustrates the spatial distributions of the H_2S and H_2O molecules in Na-BEA frameworks L-3 and L-0. In L-3, the H_2O molecules are highly localized in regions that allow for interactions with Na^+ pairs, whereas the H_2S are displaced away from Na^+ pairs into the zeolite pore interior. There is no clear preference for H_2O versus H_2S complexation for the unpaired Na^+ cations. Furthermore, the spatial distributions of the paired Na^+ cations in framework L-3 are somewhat more smeared out than those of the unpaired Na^+ cations.

The differences in the coordination of the COS and CO_2 molecules is not as pronounced (see Figure 3.4). In frameworks L-3 and L-2, the COS molecules on average orient with the O atom pointing toward the nearest Na^+ cation, but the orientation is reversed in framework L-0. Only in framework L-3 with its 3 Na^+ pairs does the coordination of O(COS) significantly exceed that of O[CO_2], whereas there is no such preference in frameworks L-2 and L-0. These structural data indicate that the enhanced conversion in Na-BEA arises predominantly from favorable interactions of H_2O with the Na^+ cations (i.e., one could also view this complexation as a secondary reaction, $Na^+ + x\ H_2O \rightleftharpoons Na(H_2O)_x$, shifting the equilibrium in Eq. 3.1 to the right in accord with Le Chatelier's principle); COS plays only a minor role in the product stabilization.

3.4 Conclusions

This study provides molecular-level insights for the selection and design of zeolites for natural gas sweetening applications that allow for better control of the undesirable conversion of H_2S and CO_2. Particularly, the present simulations show Na^+ pairs yield a larger effect on shifting the equilibrium constant than an equivalent number of isolated Na^+ cations.

RxFPMC can be applied to study a plethora of chemical reactions under confinement. Using KS-DFT to compute system energies and advanced Monte Carlo moves removes limitations imposed by pre-specified sets of reactions and ideal-gas equilibrium constants and allows to account for framework flexibility and internal sorbate structure due to strong guest–host interactions. Design of new nanoporous materials, such as zeolites and metal–organic frameworks, targeting specific applications involving chemical reactions can greatly benefit from guidelines provided by the RxFPMC simulations.

Chapter 4

Studies of Nucleation in Supersaturated Calcium Carbonate Solutions

4.1 Introduction[1]

For more than a century, nucleation from electrolyte solutions has been viewed through the lens of classical nucleation theory (CNT), which is based on two key principles [110]. First, there is a free energy barrier to nucleation due to the surface tension associated with creation of the solid–liquid interface; this leads to a critical size that clusters must exceed before the chemical potential driving force allows them to grow spontaneously. Second, clusters are able to overcome this barrier due to the inherent thermal fluctuations of the system, which allow for the addition of individual ions or

[1]This chapter represents some outcomes of a joint research project. The complete list of author contributions is available: K.H., C.B., J.M.X., S.P., J.L.F., and T.H. performed and analyzed the XANES experimental data. M.G., N.G., and J.L.F. performed and analyzed the theoretical XANES spectra. M.D.B. and C.J.M. performed and analyzed DFT and MM calculations. **E.O.F.** and J.I.S. performed and analyzed MC calculations. **E.O.F.**, J.I.S., G.K.S., B.A.L., and C.J.M. constructed and analyzed the theoretical solution model. B.A.L. and K.H. validated the solution model with known thermodynamic models. J.J.D.Y., K.H., **E.O.F.**, T.H., J.L.F., and C.J.M. wrote the manuscript. A report on this project has been published [109].

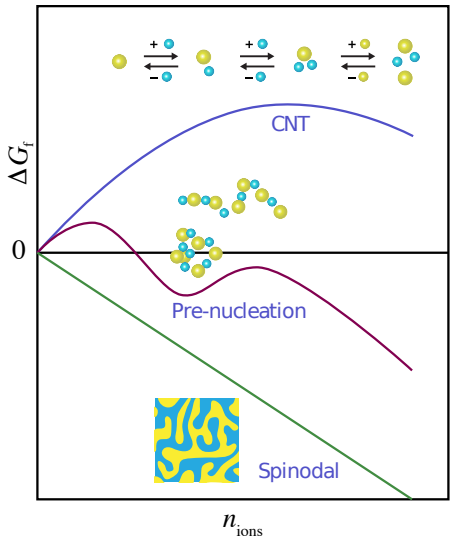

Figure 4.1: A schematic representation of the different pathways for the formation of solid polymorphs of $CaCO_3$. In CNT, nucleation proceeds uphill in free energy through the addition of Ca^{2+} (cyan), CO_3^{2-} (yellow), or ion pairs where large clusters are considered rare (but necessary) and the metastable solution phase is dominated by monomers and ion pairs. Prenucleation is characterized by a metastable solution phase dominated by populations of larger clusters that undergo further association to form the solid polymorph $CaCO_3$ [111]. Spinodal decomposition occurs spontaneously, creating the solid polymorph and ions at a high degree of supersaturation.

molecules despite the uphill trend in free energy. These two principles exhibit a self-consistency because work must be expended to increase the size of a subcritical cluster, and the number density of clusters falls off exponentially with size. Consequently, small species (that is, ions and ion pairs) dominate cluster populations (Figure 4.1). This picture of nucleation has been recently called into question based on experimental observations and classical molecular mechanics (MM) simulations of cluster formation in supersaturated solutions of $CaCO_3$, which cast doubt on one principle or on both principles [111, 112, 113, 114, 115, 116].

These studies suggest that these solutions contain populations of oligomeric species referred to as prenucleation clusters in numbers far in excess of what would be expected from classical solution models. Titration experiments demonstrate that, depending on pH, approximately 30 to 75% of Ca^{2+} ions in $CaCO_3$ solutions exist in complexes [111]. The models of nucleation based on prenucleation clusters hold that the majority of

this high degree of complexation reflects the dominance of large, oligomeric clusters [111, 112, 113]. In contrast, classical descriptions predict that species beyond ions and ion pairs should exist at concentrations that are lower by orders of magnitude [110] and thus attribute the bulk of this complexation to the formation of ion pairs, such as $CaCO_3$ (aq) or $CaHCO_3^+$ (aq), at concentrations that can be predicted on the basis of equilibrium constants for ion pairing [117]. Moreover, nucleation of the solid in prenucleation cluster models is proposed to proceed by aggregation of these oligomeric clusters with only a small free energy barrier, perhaps via the formation of a dense but disordered liquid precursor; free ions and ion pairs play little or no role in nucleation (Figure 4.1).

The essential discrepancy between CNT and prenucleation cluster models lies, then, in the free energy of cluster formation, which is directly reflected in cluster populations. However, although cryogenic transmission electron microscopy (cryo-TEM) data appeared to initially provide evidence for the existence of oligomeric clusters in $CaCO_3$ solutions [116], later experiments and reevaluation of the previous data led to the conclusion that they could not be detected by this method [118]. Thus, no oligomeric clusters of any size in saturated $CaCO_3$ solutions have been seen directly by microscopy techniques, and to date, they have evaded detection by scattering methods. Their existence was inferred from analytical ultracentrifugation (AUC) data [111, 112, 115], but no direct connection between the oligomeric species detected in AUC and the degree of complexation evident in titration experiments has been established.

Molecular simulation of solution structure is equally challenging because of the difficulties of developing pairwise additive classical MM potentials that correctly describe ion-pairing interactions and simulating sufficient periods of time to adequately sample cluster configurations. Thus, the use of standard molecular dynamics simulations of slightly supersaturated solutions to achieve equilibrium and correctly predict cluster stability, size distribution, and free energy of formation remains a formidable challenge [119] and can lead to contrasting results. Although initial applications of MM to the $CaCO_3$ system supported the conclusion that oligomeric species far larger than simple ions or ion pairs were abundant and stable relative to the latter [113], more recent work using the same classical potentials extending the simulations to longer times in dilute solutions concluded that the clusters decay into ions and ion pairs [118]. A complete

$CaCO_3$ solution model that correctly describes ion-pairing interactions, gives the distribution of species as a function of solution concentration, and correctly predicts solution equilibrium constants is lacking, as is a direct connection between the predicted species and those seen experimentally. Here, the authors combine sophisticated experiments, simulations using advanced sampling techniques and complex interaction potentials, and electronic structure theory to overcome many of these limitations and thereby achieve a detailed molecular picture of ion pairing, cluster size distribution, and the initial stages of nucleation in supersaturated $CaCO_3$ solutions. This methodology can be used to study other chemical systems as well.

4.2 Results

4.2.1 Constructing a Theoretical Solution Model

To create a theoretical solution model, we first construct potentials of mean force (PMFs) between Ca^{2+} and either CO_3^{2-} or HCO_3^- using first-principles molecular dynamics simulations with the interactions described by Kohn–Sham density functional theory (DFT) [120, 121]. The advantage of DFT over more common MM approaches is that DFT provides an excellent description of aqueous response to both small and large interfaces [122] and better representation of the local complex structure around ions, such as Ca^{2+}, as differentiated by X-ray absorption fine structure measurements [123]. Because precipitation in the $CaCO_3$ system occurs already at concentrations $< 10^{-3}$ M, accurate information on the solvent-mediated ion–ion interactions can be gleaned from simulations for systems containing a single ion pair and 96 water molecules. The comparison shown in Figure 4.2B between the PMFs of $CaCO_3$ ion pairing generated from DFT and MM representations [124, 125] illustrates the following important differences: (i) DFT yields a 50% smaller free energy gain for ion pairing (at 3 to 4 Å); (ii) DFT slightly favors bidentate versus monodentate ion pairs, whereas the opposite holds for the MM description; and (iii) DFT yields a significantly larger barrier separating contact ion pairs from solvent shared ion pairs (Figure 4.2B). The key consequence of these differences is a significantly reduced propensity to form ion pairs in the DFT model as compared to the MM model.

Knowledge of the DFT-based PMFs allows us to construct a reduced molecular

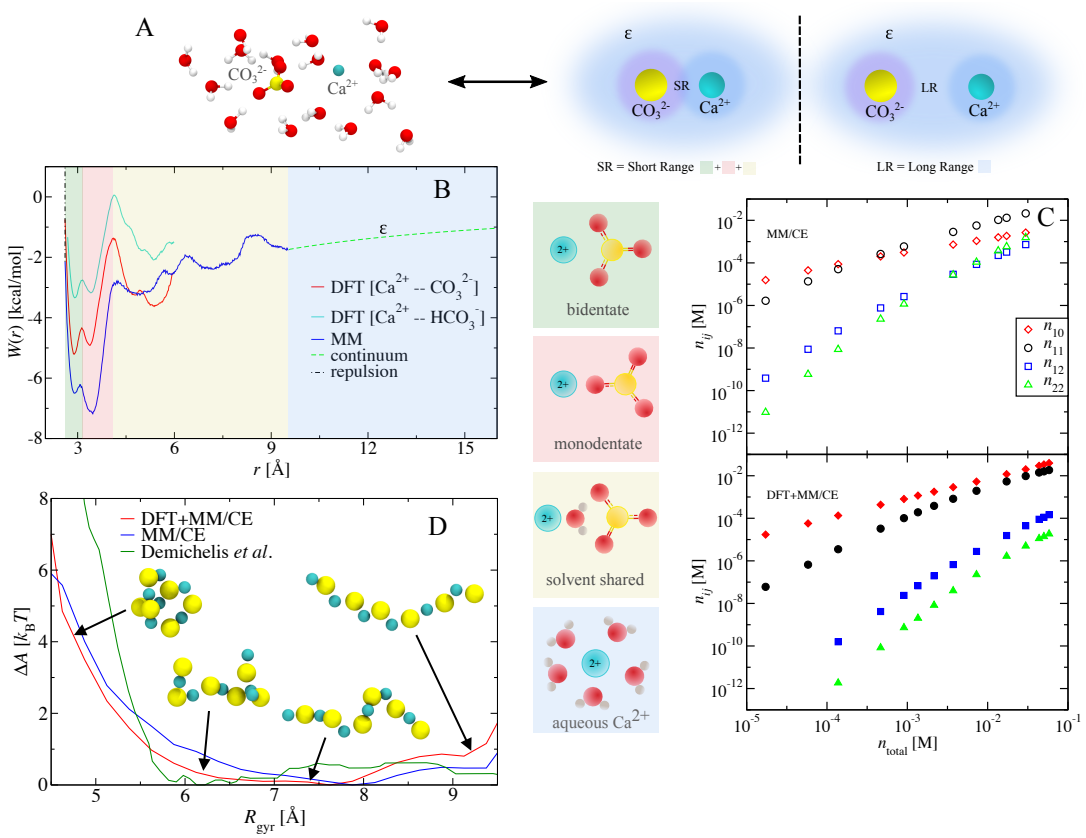

Figure 4.2: Theoretical basis of the molecular base $CaCO_3$ solution model. (A) Schematic of the mapping of the explicit molecular model to the reduced molecular model. The short-range interaction of the reduced model is computed through a PMF using a molecular model, and the long-range interaction is determined by a CE description. (B) The computed PMFs. The color-coded maps correspond to the short-range structural moieties that are depicted in the schematics. (C) Equilibrium populations n_{ij} of small $CaCO_3$ clusters versus the total concentration of $CaCO_3$ as determined by the reduced model (shown in (A), right). SEM is smaller than a symbol size. (D) Free energy of $(CaCO_3)_6$ clusters versus the radius of gyration (with some representative structures shown) comparing the present work to the explicit molecular model of Demichelis et al. [113].

representation consisting of Ca^{2+} and CO_3^{2-} ions in a solvent treated in an implicit manner (Figure 4.2A). A key assumption is that the PMFs that describe the short-range interaction are independent of concentration. For both the MM and the DFT methods, PMFs are augmented with continuum electrostatics (CE) represented by the Coulomb

potential modified by the experimentally determined dielectric constant for water to describe the long-range interactions (Figure 4.2B). Thus, we refer to the underlying potential for the reduced model for the MM as MM/CE. Because of the relatively small systems used to perform the DFT calculations, additional assumptions need to be made to construct the reduced interaction potential. Specifically, the DFT PMF requires an additional matching to the MM/CE in the solvent-shared region. This is based on the observation that the coordination number of the Ca^{2+} ion with respect to water is seven at Ca^{2+}–CO_3^{2-} distances > 4 Å reflecting the isolated cation [126]. The coordination number at this separation is also consistent with a recent two-dimensional PMF using the MM interaction potential [127]. Thus, we refer to the reduced potential based on the DFT short-range PMF as DFT + MM/CE, which is shown in Figure 4.2B. It is of interest to note that although we performed a one-dimensional PMF, we have access to Ca^{2+} coordination numbers with respect to water in our biased ensemble. For the DFT, the coordination number of Ca^{2+} with respect to water is roughly 4.5 at Ca^{2+}–CO_3^{2-} separations between 2.5 and 3.0 Å. For separations between ∼3.0 and ∼3.8 Å, the coordination number rises steeply from 4.5 to 5.5. For separations between ∼3.5 and ∼4.0 Å, the coordination number is 6. Beyond the distance of ∼4.0 Å, the coordination number is 7. Last, the like-charge repulsive interactions for both the DFT + MM/CM and the MM/CM models were taken from the PMFs using the MM model.

We use this reduced model to predict cluster size and shape distributions and the equilibrium constants between the interconversion of species as a function of concentration. Recent work suggests that cluster configurational entropy can significantly influence the size and shape distributions of small nucleation clusters [128], making efficient statistical sampling of the molecular configurations for concentrated $CaCO_3$ a necessity. In keeping with the assumption of a dilute limit, we neglect many-body correlations between the ions but retain both two-body correlation effects and the important water-ion correlations that are encoded in the PMFs (Figure 4.2B), as done previously in predicting the collective phenomena of electrolytes at low concentrations [126, 129]. Thus, our reduced model comprises two different attractive point particles to represent Ca^{2+} and CO_3^{2-} (Figure 4.2A) that interact according to the corresponding PMFs (Figure 4.2B), where the corresponding terms for Ca^{2+}–Ca^{2+} and CO_3^{2-}–CO_3^{2-} are strictly repulsive.

This reduced model in conjunction with the aggregation-volume-bias Monte Carlo (AVBMC) method (see Computational Methods for details) [13] allows for the fully converged statistical sampling required to obtain the cluster concentrations, n_{ij}:

$$n_{ij} = \exp(-\beta G_{ij})(n_{10})^i(n_{01})^j. \tag{4.1}$$

The subscript ij denotes stoichiometric Ca^{2+} and CO_3^{2-}, respectively, $\beta = 1/k_BT$ (where k_BT is Boltzmann constant multiplied by absolute temperature), and G_{ij} is the free energy of formation of the ij cluster relative to the corresponding nonassociated cations and anions. The resulting cluster populations for the DFT + MM/CE description are dominated by free ions (n_{10} and n_{01}) and ion pairs (n_{11}), and the concentrations of all larger species are orders of magnitude smaller (Figure 4.2C, bottom), for example, by > 100 and > 1000 for n_{21} and n_{22}, respectively.

Further connection of our MM/CE version of our solution model to previous MM simulations that used explicit water is achieved by directly comparing shape distributions for $(CaCO_3)_6$ clusters to that of Demichelis et al. [113] (Figure 4.2D). This comparison yields good agreement, demonstrating the efficacy and dominance of the PMF in describing ion pairing under dilute solution conditions. An important prediction of our simulations is that, despite the agreement in cluster shape distribution obtained here with MM/CE, DFT + MM/CE, and past simulation studies [113], neither our MM/CE nor our DFT + MM/CE models find larger species, such as $(CaCO_3)_6$, to be present in solution at any relevant concentrations. From the fact that we are able to converge the sampling of the cluster populations at the correct experimental concentrations, we conclude that these larger species are insignificant in slightly and moderately supersaturated solutions.

4.2.2 From PMFs to Bulk Thermodynamic Properties

Taking the species populations from the solution model, we calculate the equilibrium constants $K_{ij} = n_{ij}/(n_{10}^i n_{01}^j)$ for interconversion between species (Figure 4.3B). We use these DFT-based constants to model the free Ca^{2+} concentration in bulk solutions during titration of Ca^{2+} into either carbonate or bicarbonate buffers and obtain good agreement between the predictions and previously reported experimental titration data

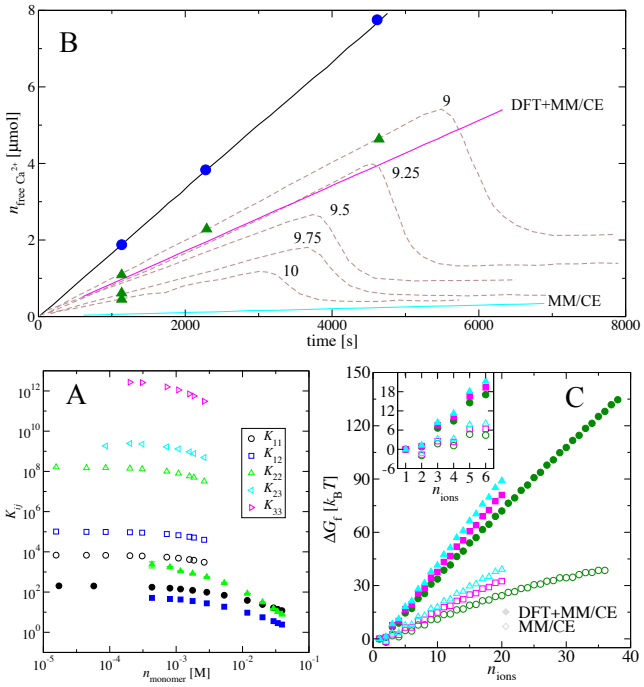

Figure 4.3: Macroscopic outcomes of the solution model. (A) Computed (magenta and cyan) and experimental [111] titration curves of $CaCl_2$ into a carbonate buffer solution at different pH values: The black solid line shows the total amount of Ca^{2+} titrated into solution, and brown dashed lines show the amount of unbound Ca^{2+}. The blue circles correspond to the total amount of calcium added to solutions and the green triangles refer to the amount of free calcium in the solution under conditions where the XANES experiments are conducted. (B) Computed equilibrium constants for the formation of small $CaCO_3$ clusters versus equilibrium concentration of monomers. Filled symbols denote DFT + MM/CE results, and open symbols indicate MM/CE results. SEM is smaller than a symbol size if not shown. (C) Free energy of formation of $CaCO_3$ clusters versus cluster size. Dark green, magenta, and cyan symbols correspond to total $CaCO_3$ concentrations of 17.22, 7.264, and 3.719 mM, respectively. These results represent the free energy landscape for nucleation and are distinct from those of Wallace et al. [114], which only consider the free energy of individual clusters. SEM is smaller than a symbol size.

(Figure 4.3A) [111]. Although this calculation ignores the secondary effects of the background electrolyte, Na^+ and Cl^-, inclusion of these species in classical solution models demonstrates that these species have negligible contribution to the total speciation (see Table S3 in the original reference [109]). Moreover, classical solution models based on

accepted equilibrium constants [117] predict amounts of ion pairs and titration curves similar to those obtained in the DFT + MM/CE model. In contrast, because of the stronger ion pairing seen in the MM model that leads to 10- to 100-fold increases in predicted ion complexation (Figures 4.2C and 4.3B), the degree of ion complexation occurring during Ca^{2+} titration is strongly overestimated (Figure 4.3A). Although both the DFT + MM/CE and MM/CE models produce the same qualitative conclusions regarding speciation, the bulk titration data support the picture put forth by the DFT-based PMFs where free Ca^{2+} followed by ion pairs, with almost negligible contribution from larger clusters, dominate at experimentally relevant concentrations.

Using umbrella sampling [19] in conjunction with AVBMC, we also calculated the free energy of cluster formation ΔG_{f} as a function of cluster size (Figure 4.3C). In keeping with the predicted dominance of ions and ion pairs, we find that ΔG_{f} exhibits the size dependence expected from CNT using either the DFT + MM/CE or MM/CE solution models, but the greater tendency toward ion pairing in the MM/CE model leads to a weaker ΔG_{f} dependence and thus smaller nucleation barriers. Using the ΔG_{f} size dependence of the DFT + MM/CE model, we deduce that the interfacial free energy for the formation of amorphous $CaCO_3$ (ACC) is between 14 and 25 $\mathrm{mJ/m^2}$, depending on the number density [130]. This value implies that at a supersaturation of between 1.5 and 2.0, the barrier to ACC formation falls below $k_{\mathrm{B}}T$ and the solution spontaneously phase-separates. Assuming solubility on the order of 10 to 15 mM for ACC, this estimate compares favorably to the recent experimental determinations of the point at which $CaCO_3$ solutions undergo spinodal decomposition [131].

Following Wallace et al. [114], an important self-consistent check of our reduced model is to convert the ideal solution distribution coefficients into Gibbs free energies of cluster formation, G_{ij}, using the absolute energy scale:

$$K_{ij} = e^{-k_{\mathrm{B}}TG_{ij}} \frac{\lambda_{ij}}{\lambda_{10}^{i}(\lambda_{01}q_{\mathrm{rot}})^{j}}; \tag{4.2}$$

$$\lambda_{ij} = \left[\frac{(im_{\mathrm{Ca}} + jm_{\mathrm{CO_3}})k_{\mathrm{B}}T}{2\pi\hbar^2}\right]^{\frac{3}{2}}; \tag{4.3}$$

$$q_{\mathrm{rot}} = \sqrt{\pi I_a I_b I_c}\left(\frac{2k_{\mathrm{B}}T}{\hbar^2}\right)^{\frac{3}{2}}. \tag{4.4}$$

Here, λ_{ij} is the thermal de Broglie wavelength, and q_{rot} is the rotational partition function. Please note that, in accord with Ref. [114], the vibrational partition function of the CO_3^{2-} anion is not considered and only a reduced rovibrational partition function of the aggregates is included when sampling considers the CO_3^{2-} anion as a point particle. k_{B}, T, and $\hbar$ are the Boltzmann constant, the absolute temperature, and the reduced Plancks constant, respectively. m_{Ca} and $m_{\mathrm{CO_3}}$ are the masses of a Ca atom and CO_3^{2-} compound, and $I_{a,b,c}$ are the components of the moment of inertia of CO_3^{2-}. Here, we model CO_3^{2-} as a planar compound with three-fold symmetry. Using the K_{ij} as defined above we can obtain the free energies of cluster formation and directly compare the reduced model based on the PMFs of ion-pairing (see Figures 4.3B and 4.3C) to the values obtained by Wallace et al. [114] using the all-atom interaction potential in the presence of explicit water [124, 125]. The results are shown in Figure 4.4. One discrepancy should be emphasized however. Originally reported values of G_{ij} were confused with free energies of formation and supersaturated $CaCO_3$ solutions were believed to spinodally decompose due to monotonically decreasing free energies. However, there is a clear difference between absolute free energies and free energies of formation, which has to be referenced to a monomer chemical potentials as well as supersaturation. Thus, we report both free energies (Figure 4.3C and Figure 4.4). Our reduced model for the MM PMF is in excellent agreement for $n_{i+j} \leq 25$ with the cluster free energies obtained by Wallace et al. [114]. Taken together with the data shown in Figure 4.2D of the main text based on the study of Demichelis et al. [113], it is clearly evident that the reduced model can reproduce both the cluster free energies and the populations of structural moieties using the all-atom empirical force field. Deviations from the data of Wallace et al. [114] at larger cluster sizes are likely due to under converged sampling in the molecular dynamics simulations and the underlying approximate methods (neglect of the rotational partition function of the clusters) used to compute the cluster free energies in Ref. [114]. To demonstrate the excellent agreement obtained from AVBMC simulations (and, hence, sufficient sampling) at three different concentrations, Figure 4.4 shows the differences in the cluster free energies, which are extremely small between low and intermediate concentrations and on average less than 1% between intermediate and high concentrations, where non-idealities become more important.

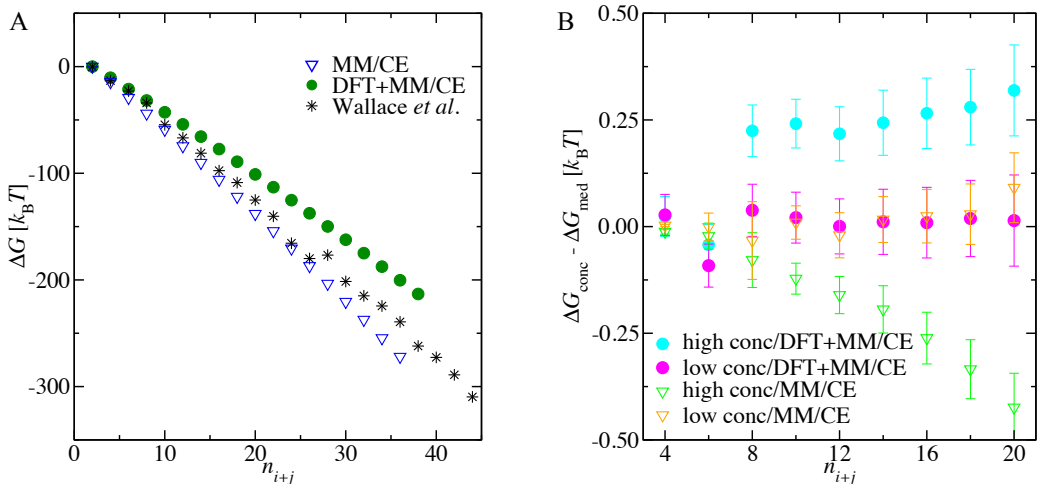

Figure 4.4: Summary of the calculated free energies. (A) Comparison of the cluster size free energies referenced to G_{11} as defined in the study of Wallace et al. [114]. Only the simulation data at the highest density are used for this graph. (B) Differences in computed cluster free energies at the low and high concentrations with respect to the median concentration. Error bars correspond to 95% confidence intervals.

4.2.3 In situ molecular-level measurements of solution speciation[2]

The Ca speciation in $CaCO_3$ solutions was experimentally determined in situ by using synchrotron-based Ca K-edge X-ray absorption near-edge structure (XANES) spectroscopy [132, 133] to directly probe the local chemical environment of aqueous Ca^{2+} ions. Spectra were taken using a liquid microjet [134, 135] at specific Ca^{2+} concentrations and pH values lying along published titration curves (Figure 4.3A, blue circles and green triangles) [111]. The use of state-of-the-art undulator beamlines [136] enabled measurements on dilute supersaturated solutions extending down to submillimolar Ca^{2+} concentrations that are close to the solubility limit of the most stable phase (calcite). XANES spectra for various solid carbonates were also acquired for comparison.

The obtained XANES spectra of both the Ca^{2+} (for example, $CaCl_2$) and $CaCO_3$ solutions (Figure 4.5A) are distinct from one another and from those of solid ACC,

[2]All experiments and TDDFT calculations were not performed by the author (E. O. Fetisov). However, this section builds upon results from the MC simulations and connects those results to experimental findings. For a much more detailed explanation of the experiments the reader is referred to the original publication [109].

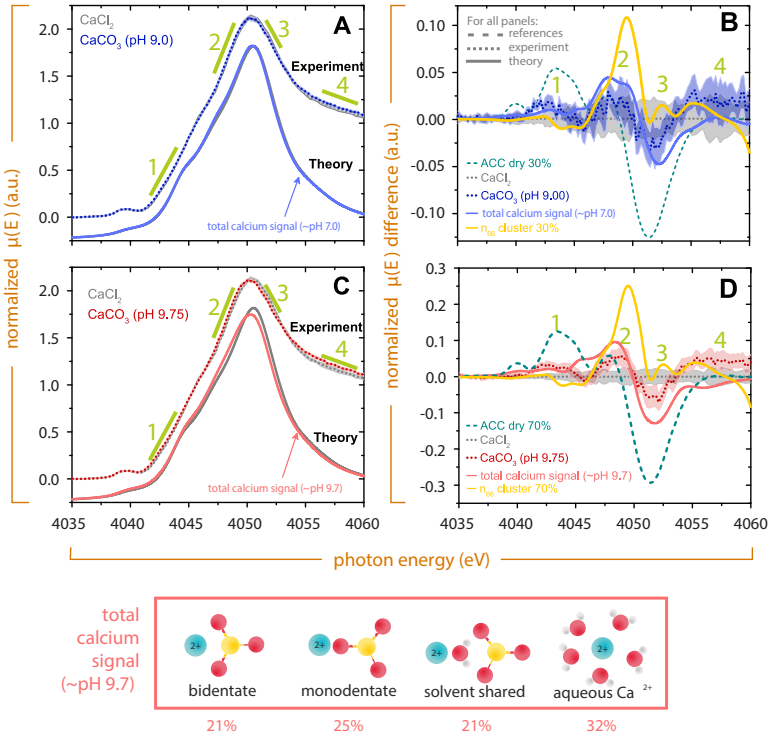

Figure 4.5: Comparison of experimental and theoretical XANES spectra of carbonate solutions. (A) Experimental XANES spectra for a pure $CaCl_2$ solution (gray dots) and for a freshly mixed 1:1 solution of $CaCl_2$ (150 μM; pH 9.0) and carbonate buffer (20 mM; pH 9.0) (blue dots), compared to the simulated spectra for a $CaCl_2$ (as Ca^{2+}) solution (gray line) and for a mixture of $CaHCO_3^+$ ion pairs with speciation as calculated from DFT theory (blue line). The regions marked as 1, 2, 3, and 4 indicate the regions where the spectrum of a Ca^{2+} solution differs from a solution with ion pairs. (B) Difference spectrum between a solution with ion pairs and the $CaCl_2$ standard solution for the cases shown in (A). The blue dotted line with a light blue error region represents the difference spectrum between the $CaCO_3$ and $CaCl_2$ solutions of the experimental data shown in (A). The solid blue line represents the difference spectrum for the theoretical data shown in (A). For comparison, the difference spectra of the experimental ACC (dry) and the theoretical $(CaCO_3)_6$ clusters are shown. For these two standards, a mixing ratio was applied to match the amount of bound species measured by Gebauer et al. [111], in which 30% is the amount of bound species in a supersaturated $CaCO_3$ solution at pH 9.0. (C and D) Same set of analysis for a solution at pH 9.75, where the speciation percentages are determined from the R^2-weighted integral of the DFT + MM/CE potential in Figure 4.2B.

aragonite, and calcite in the regions of the pre-edge, near-edge, and white line (see Figure S6 in the original reference [109]). The normalized difference spectra (Figure 4.5B), that is, the difference between the $CaCO_3$ solution spectra and the spectrum of fully hydrated Ca^{2+} (in a $CaCl_2$ solution), highlight the spectral changes that are due to the presence of associating ions whose structure deviates from that of normal sevenfold coordinated aqueous Ca^{2+} ions [137]. It was observed that, for constant solution pH, the difference spectra show little change with Ca^{2+} concentration, whereas with increasing pH (9 to 10), the magnitude increases (see Figure S2 in the original reference [109]). The possibility of compact nanoclusters consisting of > 10 cation/anion assemblies that are structural analogs of ACC, calcite, or aragonite having XANES spectra that are similar to those of the bulk solid [138] was also considered. A linear combination of any of these solid phases with a representative fraction of Ca^{2+} yielded a unique concentration-dependent XANES spectrum that does not reproduce the experimental $CaCO_3$ solution spectra. Thus, one can exclude the presence of nanoscale solid $CaCO_3$ phases [119] even at supersaturations substantially above the phase boundary seen in the titration data (Figure 4.3A).

To identify the ion-pair species that are responsible for the observed difference spectra, the time-dependent DFT (TDDFT) XANES spectra [139, 140, 141] were computed using the species and distributions harvested from the DFT-based PMFs and the AVBMC populations in Figure 4.2 (B and C, respectively). When the TDDFT method was applied to the solid-phase standards, the calculated and experimental spectra were in reasonable agreement (see Figure S6 in the original reference [109]). For the aqueous phases, the DFT methods were used to calculate the spectra for mixtures of monodentate and bidentate contact ion pairs, solvent-shared ion pairs, and free Ca^{2+} in the two limiting conditions of pure carbonate (pH 9.75) and pure bicarbonate (pH 7.0). These two theory conditions bracket the experimental pH values, providing an approximate amplitude estimate for the experimental spectra acquired at pH 9.0 and 9.75. Intermediate pH values would be represented as a linear combination of these carbonate, bicarbonate, and free Ca^{2+} theory spectra. The perturbation of the aqueous Ca^{2+} structure due to the presence of the anion in the first or second solvation shell alters the XANES spectra in different ways. For the experimental condition of pH 9.75 in Figure 4.5 (C and D), the theoretical distribution of Ca species (Ca^{2+} with CO_3^{2-}) is

in agreement with experimental models based on available thermodynamic data. The predicted difference spectrum (Fig 4.5 D) shows exceptional agreement with the experimental spectrum in both the position of the features and their amplitudes that lie just outside of the experimental uncertainties.

Although our MM- and DFT-based solution models predict only minute concentrations of multi-ion clusters (for example, $(CaCO_3)_6$), it is instructive to ascertain whether their XANES features can be used to rule them out. TDDFT XANES spectra were computed for a representative mixture of the four $(CaCO_3)_6$ clusters ranging from globular to linear (Figure 4.2D) that were predicted by the MM model at a concentration needed to account for the bound species measured experimentally [111]. The resulting spectrum exhibits a line shape that is not consistent with the experimental solution spectra (see Figure S6 in the original reference [109]). Thus, the ability to quantitatively calculate the XANES spectrum of candidate Ca^{2+} cluster species provides the opportunity to test their existence as components in experimental spectra, and the results obtained here imply that they are not present in significant numbers.

4.3 Discussion

The agreement between the experimental and calculated XANES spectra demonstrates the exquisite sensitivity of the local electronic response to changes in the structural motifs in the first and second solvation shells of Ca^{2+} ions. The agreement between the XANES difference spectra under varying pH conditions strongly supports the conclusion that the solution structure is dominated by ion pairing, even near the solution–solid-phase boundary. The observed trends in both the experimental and theoretical XANES spectra are also consistent with speciation calculations based on the previously determined and widely used experimental equilibrium constants [117]. These equilibrium constants corroborate the dominance of ions and ion pairs and show that the overall fraction of "bound" n_{11} $CaCO_3$ clusters increases with pH at constant concentrations, whereas the ratio of free $CaCO_3$/bound (Ca) remains constant during titration at constant pH. Both observations are consistent with the titration experiments (Figure 4.3B).

Although the calculated free energy of cluster formation (Figure 4.3C) predicts that large clusters are necessary for nucleation, the speciation analysis, both theoretical and

experimental, implies that these clusters are exceedingly rare except at supersaturations for which the free energy barrier approaches $k_{\mathrm{B}}T$. Thus, our findings rule out the presence of large prenucleation clusters as a dominant species, producing a clear picture of $CaCO_3$ nucleation proceeding by monomer addition from a solution rich in isolated ions and their pairs, although our findings are agnostic with regard to whether the first condensed phase is solid ACC or a dense liquid state forming in a binodal region as concluded in previous computational [114] and experimental [118] studies. This picture is further corroborated by XANES experiments demonstrating that the speciation is unchanged as a function of concentration at a fixed pH and ion pairing becoming less prominent at lower pH, both in accordance with a linear titration curve (Figure 4.3A), where the ratio of bound/unbound Ca depends solely on pH. Moreover, it is in agreement with the conclusions of Smeets et al. [118], in which cryo-TEM was used to rule out the existence of clusters larger than 9 Å in diameter and MM simulations were extended to adequate time scales to predict that clusters placed in dilute solutions decay into ions and ion pairs and solutions containing free ions generated few additional species beyond ion pairs [118].

It is worthwhile to briefly mention two equally important criteria that must be fulfilled to meet the most stringent requirements of a particular speciation model. First, the experimental titration curve is linear, reflecting equilibrium with simple ion pairs such as $CaCO_3$. In contrast, a model invoking equilibrium between free ions with primarily a dominance of larger, oligomeric clusters would be strongly nonlinear in accordance with Equation 4.1. Second, an examination of the titration data by Gebauer et al. [111] reveals that, despite significant changes in total calcium concentration, the total amount of bound calcium at the moment when the nucleation starts is a constant of $\sim$3.9 μM for all pH values in line with observations of Kellermeier et al. [142], adding additional credence to the validity of the solution model derived herein that ion pairs are the species responsible for cluster formation and eventual precipitation. Hence, it is their concentration that determines the onset of nucleation. Our conclusion that ion pairs are the fundamental unit leading to nucleation of the first $CaCO_3$ condensed phase parallels recent findings that this same species is the fundamental unit leading to growth of calcite [143, 144, 145].

There is close agreement between the predicted and published equilibrium constants,

titration curves, XANES spectra, and pH dependence of bound Ca^{2+}. There is agreement with the classical dependence of the free energy of cluster formation on size and between the predicted and measured concentrations at which spontaneous phase separation occurs. In total, $CaCO_3$ solutions contain negligible concentrations of species beyond simple ions and ion pairs and form clusters at a rate, in relative numbers, and with energies of formation that are expected fromCNT. Hence, all these factors point to a picture of supersaturated $CaCO_3$ solutions defined by classical views of solution speciation, cluster populations, and the initial stages of nucleation.

4.4 Computational Details

4.4.1 Construction of Interaction Potentials

Born–Oppenheimer first-principles molecular dynamics simulations in the canonical ensemble NVT (at $T = 300$ K) using periodic boundary conditions were performed using the CP2K simulation suite containing the Quickstep module for the density functional theory (DFT) calculations [48, 52]. A protocol similar to that of Martyna et al. [146] was followed using a double-zeta basis set that has been optimized for the condensed phase [52] in conjunction with Goedecker–Teter–Hutter pseudopotentials [147] using a 400-rydberg cutoff for the auxiliary planewave basis. A Nosé–Hoover chain thermostat was attached to every degree of freedom (translations of nuclei) to control the temperature [146]. The Becke exchange [49] and correlation due to LYP (Lee–Yang–Parr) [148] was used in addition to the dispersion correction (D2) put forth by Grimme [149] with a 40 Å cutoff (that is, beyond the minimum image convention).

The carbonate/bicarbonate simulation cell consisted of one carbonate anion in 95 water molecules and a single Ca^{2+}, in a cubic box with an edge length of 14.40 Å. The PMF was constructed by restraining the carbon of the CO_3^{2-} or HCO_3^- species and the Ca^{2+}, where sampling windows over the distance ranging from 2.4 to 5.6 Å were equally spaced by 0.2 Å using harmonic umbrella potentials of the form $V_{\text{umbrella}} = k(r - r_0)^2$ with a force constant k of 4000 kJ mol^{-1} nm^{-2}. To ensure sufficient sampling in the barrier region, additional windows with a stiffer force constant were added, ranging from 3.6 to 4.6 Å equally spaced apart by 0.1 Å with a force constant of 14,000 kJ mol^{-1} nm^{-2}. In each umbrella window, a trajectory of at least 40 ps was collected after

5 ps of equilibration. The weighted histogram analysis method was used to extract a free energy profile from these histograms [26]. Simulations for the MM potential were performed using the interaction potential of Raiteri and co-workers [124, 125] under system size and simulation protocols identical with those used for the DFT calculation, but longer simulation times (1.5 ns per window) were accessible for the MM simulations. To ensure that there were no significant finite size effects, PMF for the MM potential was also calculated using a larger simulation cell with linear dimensions of 28 Å.

In principle, the above protocol should converge to a proper PMF that is correctly weighted with respect to orientation. To directly test understanding of how the PMF can be affected by the nonspherical nature of the carbonate anion, additional calculations were performed. To this end, PMFs (both DFT and MM) were constructed using the same protocol as outlined above by restraining the Ca^{2+} (i) along the bisector of a single O–C–O bend of the carbonate (bidentate), (ii) along the C–O of the carbonate, and (iii) perpendicular to the plane of the carbonate molecule. The results corroborate the understanding that much of the attractive interaction is in the plane of the carbonate anion. Unfortunately, it is not a straightforward task to simply add the three individual obtained PMFs to construct the ones used in Figure 4.2B. This would require the proper weighting of the volume of phase space (configuration space) for each of the PMFs or an interpolation of the PMFs for other sets of orientations and, in addition, the setting of an absolute scale. Nevertheless, using the three individual PMFs as a basis, it is possible to associate features of the PMFs shown in Figure 4.2B used in this study with specific structural moieties, such as monodentate and bidentate. Moreover, the overall agreement with the $(CaCO_3)_6$ structures and populations with the work of Demichelis et al. [113] demonstrates that the potentials in Figure 4.2B are general enough to reproduce subtle structural features.

4.4.2 Monte Carlo Simulations

Monte Carlo (MC) simulations were performed in the canonical ensemble at $T = 298$ K and for simulation volumes of 8×10^3 nm$^3 \leq V \leq 27 \times 10^6$ nm^3 [150]. All systems consisted of 300 Ca^{2+} cations and 300 CO_3^{2-} anions, where each ion is represented as a single interaction site. The potential energy of a configuration was calculated as the pairwise-additive sum of distance-dependent Ca^{2+}/CO_3^{2-}, Ca^{2+}/Ca^{2+}, and CO_3^{2-}/CO_3^{2-} PMFs

that are truncated at 20 Å. To efficiently sample ion addition or removal to and from clusters, AVBMC moves [13] were used in addition to standard translational moves; the ratio between attempted translational and AVBMC moves was set to 4:1. Maximum displacements of translational moves were adjusted during equilibration to yield an acceptance ratio of 40%; AVBMC moves used 2.7 Å $\leq r \leq$ 3.9 Å to define the bonding region and explored 32 trial locations, which resulted in an acceptance rate of about 9%. In the current simulations, a distance-based criterion is used to define the ion–ion aggregate size distribution. To this end, an ion belongs to a given aggregate if its distance to at least one other unlike ion in the aggregate is less than 4.0 Å. Two different sets of simulations were carried out: On one hand, simulations without additional umbrella potentials were used to determine the size (numbers of cations and anions) and shape distributions of small aggregates. On the other hand, simulations with an additional umbrella potential acting on only one specific aggregate were used to probe the size and shape distributions of larger aggregates along the nucleation pathway. For the former set, 16 independent runs were carried out for each concentration. At least 200,000 MC cycles (where a cycle consists of $N = 600$ randomly selected MC moves) were used for equilibration, and at least 1,000,000 MC cycles were used for the production periods. In the other case, one specific ion was tagged to define the aggregate, and self-adaptive umbrella sampling [19] was used to flatten the size distribution for this particular cluster, thereby providing access to reliable statistics for much larger aggregates. Eight independent runs were carried out consisting of at least 200,000 MC cycles for the refinement of the umbrella potential, followed by at least 700,000 MC cycles for production.

Chapter 5

First Principles Molecular Dynamics Study of a Deep Eutectic Solvent: Choline Chloride/Urea and Its Mixture with Water

5.1 Introduction[1]

Ionic liquids (ILs), generally defined as solvents entirely comprised of ions, are a topic of scientific interest due to their unique physiochemical properties and their potential to be highly tunable for a wide range of industrial applications [152, 153, 154, 155, 156]. Deep eutectic solvents (DESs), exhibiting a melting point much lower than the melting points of their constituents, are a related class of solvents usually made from a salt with a cation capable of forming H-bonds and a neutral H-bond donor molecule. Like ILs, DESs have a very low vapor pressure, are non-flammable, and offer tunability [157]. In

[1]This chapter describes the outcome of a research project carried out by E. O. Fetisov in collaboration with D. B. Harwood, I-F. W. Kuo, S. E. E. Warrag, M. C. Kroon, C. J. Peters and advised by J. I. Siepmann. A report on this project has been published [151]. Reprinted with permission from [151]. Copyright 2018 American Chemical Society.

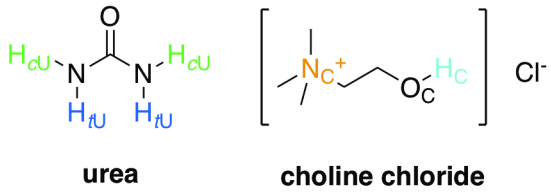

Figure 5.1: Structural formula for reline highlighting the two types of hydrogen atoms in urea and the nitrogen, oxygen, and hydroxyl hydrogen in choline that are important for its liquid structure.

2003, Abbott et al. [158] introduced the 1:2 choline chloride/urea mixture, sometimes called reline (see Figure 5.1), as the archetypical DES. Compared to ILs, DESs may offer several advantages—simple preparation and low-cost synthesis, chemical inertness with respect to water, and environmental friendliness (biodegradable and non-toxic)—which make them an attractive substitute for ILs in many applications [157, 159, 160].

Despite being a relatively new field, DES applications ranging from metallurgy to fine synthesis have emerged. For example, DESs have been used in the synthesis of various inorganic materials [161], metal coatings [162, 163], in the purification or production of biodiesel [164], and as water-immiscible extraction solvents [165]. Valuable work has been done characterizing some DESs while also confirming their potential to be tuned via the addition of water or by varying composition [166, 167].

The formation of strong H-bonds between the anion and the various H-bond donors is clearly a key feature for the suppression of the melting point of DESs [158, 168], and only recently was the liquid structure of choline chloride/urea investigated by neutron diffraction experiments with empirical potential structure refinement [169]. From a theoretical standpoint, DESs present a challenge due to polarization, charge transfer [170], and the long relaxation times associated with the formation of strong H-bonds [159]. Previous theoretical studies have mostly focused on the choline chloride/urea DES, which exhibits a large melting point depression at the 1:2 ratio of its components. Liquid reline has been probed previously using molecular dynamics (MD) simulations with the AMBER force field by Perkins et al. [171, 172] and with a modified version of the OPLS-AA force field by Sun et al. [173], and Shah and Mjalli [174] employed the MMFF (urea and choline) and SPE (water) force fields to study different mixtures of the DES with water. These MD simulations using non-polarizable force fields yielded reasonable

agreement with experimental measurements of specific densities, diffusion coefficients, and viscosities, but required *ad hoc* tuning of the partial charges. Furthermore, Garcia et al. [175] assessed ten different schemes for the assignment of partial charges for MD simulations of the choline chloride/levulinic acid (1:2) DES and reported large differences in the thermophysical properties. Similarly, Mainberger et al. [176] investigated choline chloride/(levulinic acid or glycerol or 1,4-butanediol) and reported significant differences in structural and thermophysical properties (e.g., changes in the diffusion coefficients by a factor of 20) between two different force fields or when the partial charges are scaled by 75%.

Given the importance of polarization and charge transfer for the formation of strong H-bonds in DESs, this study utilizes first principles molecular dynamics (FPMD) simulations to investigate the structure of anhydrous reline (AnR) and its equimolar mixture with water. Recently, Zahn et al. [170] used FPMD simulations to explore charge spreading and H-bonding in the choline chloride/(urea or ethylene glycol or oxalic acid) systems. In related studies, the FPMD technique has been extensively used to provide insight into structural features, dynamics, electronic properties, vibrational spectra, solvation, and reactions in various ILs [177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188].

5.2 Computational Details

The choline chloride/urea (1:2) DES and its equimolar mixture with water are studied at two different temperatures (333 K and 363 K) and the corresponding experimental densities [189]. The FPMD simulations are carried out for systems that contained 8 choline chloride ion pairs and 16 urea molecules for AnR, and an additional 24 water molecules for hydrous reline (HyR), for a total of 304 and 376 atoms, respectively. The corresponding linear dimensions for the cubic simulation boxes are: 14.311 and 14.401 Å for AnR at $T = 333$ and 363 K, respectively, and 15.404 and 15.500 Å for HyR at $T = 333$ and 363 K, respectively. The systems are pre-equilibrated via Monte Carlo simulations using only rigid-body translations and rotations and the interactions described by the modified AMBER (GAFF) force field for choline chloride and urea [171] and the TIP4P force field [190] for water. These initial equilibration periods consist

of 320,000 and 560,000 randomly selected Monte Carlo moves for the AnR and HyR systems, respectively.

All FPMD simulations are performed with the *CP2K* simulation package [96], which solves the Kohn–Sham formulation of density functional theory with the Gaussian plane wave method as implemented in the *Quickstep* module [48]. The Becke–Lee–Yang–Parr (BLYP) functional [49, 50] with the third-generation dispersion correction (D3) of Grimme [98] is used along with the triple-zeta, double polarization basis set [52], Goedecker–Teter–Hutter (GTH) pseudopotentials [53, 54], and a plane wave cut-off of 400 Ry. (For comparison, Zahn et al. [170] used a larger system size for AnR (18 choline chloride ion pairs), the D2 dispersion corrections, a molecularly optimized double-zeta basis set, and a plane wave cut-off of 280 Ry.)

Here, the time step is chosen to be 0.5 fs, and the temperature is maintained with Nosé–Hoover chain thermostats [191, 192, 146]. After geometry optimizations, the systems are equilibrated for 40 ps, followed by production periods of at least 100 ps (111 and 108 ps for AnR at 333 and 363 K, respectively; 104 ps for HyR at both temperatures). Data for the properties reported in this paper are collected every ten time steps during the production runs. To provide an estimate of the statistical uncertainties, the production periods are divided into four blocks and the corresponding standard deviations are reported here unless otherwise noted.

5.3 Results and Discussion

5.3.1 Radial Distribution Functions and Coordination Numbers

The liquid structure of the AnR and HyR systems is discussed here in terms of center-of-mass and atom–atom radial distribution functions (RDFs), coordination numbers, and an analysis of the H-bond network. The center-of-mass RDFs (see Figure 5.2) can provide insight into the overarching structure of the AnR and HyR systems. However, the complex nature of urea and choline with multiple H-bonding sites leads to broad peaks and poorly defined minima. Agreement with the RDFs for AnR obtained from neutron diffraction experiments with empirical potential structure refinement (ND/EPSR) [169] is very good with regards to peak positions (see Table 5.1) and satisfactory with regards to peak heights and shapes. It should be noted that the EPSR approach leads to larger

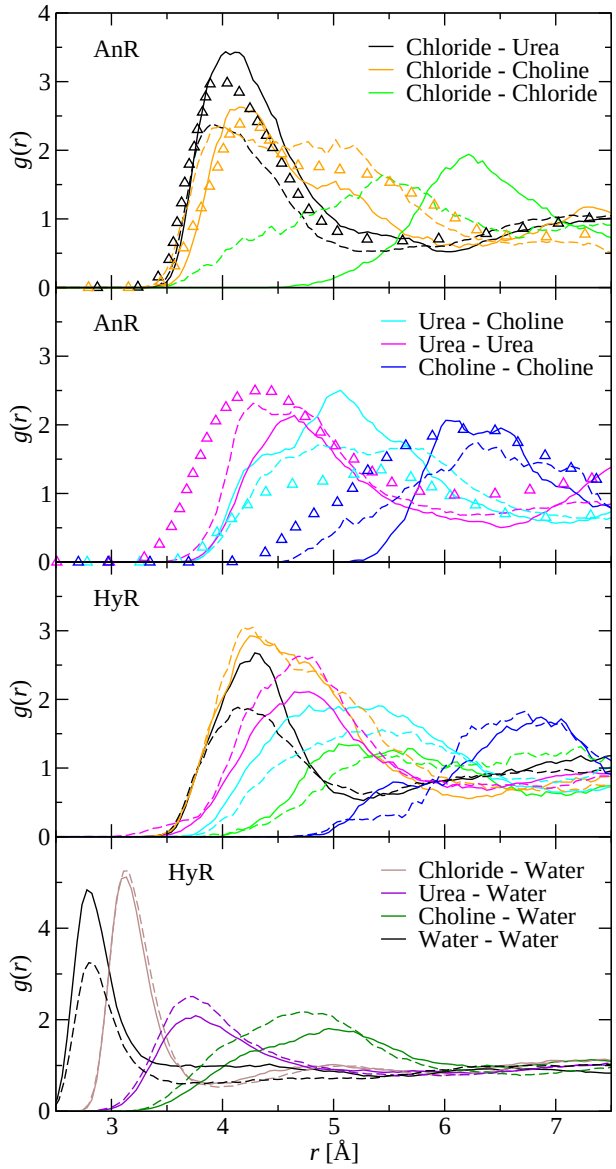

Figure 5.2: Center-of-mass RDFs for AnR (top two graphs) and HyR (bottom two graphs). The solid and dashed lines show the FPMD simulations at $T = 333$ and 363 K, respectively, and the triangles denote the ND/EPSR data (303 K) [169].

uncertainties for center-of-mass RDFs of relatively large compounds, such as urea and choline. The chloride–urea peak from the FPMD simulations is centered at 4.1 Å and relatively sharp, whereas the chloride–choline RDF exhibits a double peak at 4.2 and

5.0 Å, i.e., indicating two binding motifs for this pair. The urea–urea peak at 4.6 Å is fairly broad, and its onset is shifted by 0.3 Å to shorter separations for the ND/EPSR data. The choline–choline RDF shows a single, but very broad peak centered at 6.3 Å, whereas the urea–choline RDF exhibits a broad feature consisting of multiple merged peaks 4.4, 5.0, and 5.7 Å. It should also be noted that normalization of the RDFs leads to a noticeable suppression of the like–like RDFs for the FPMD simulations due to the small system size, e.g., pair distances are only computed for 7 choline particles surrounding a central choline particle, whereas the density is based on all 8 choline particles. As one should expect, only small changes in the RDFs obtained from the FPMD simulations are observed as temperature is increased from 333 to 363 K, but the slightly broader RDFs obtained at 363 K for urea–urea, urea–choline, and choline–choline better match the ND/EPSR data obtained at 303 K.

The corresponding coordination numbers for the first shell, N_{coord}, are listed in Table 5.1. The results from ND/EPSR and FPMD simulations agree within statistical uncertainties, but it should be noted that the FPMD simulations yield N_{coord} values that are smaller by 20-30% for choline–chloride, chloride–urea, and urea–urea pairs, whereas N_{coord} for the choline–urea pair is larger by a factor of 1.4.

Table 5.1: Positions (in Å) of First Maximum and Minimum in Selected RDFs and Average Coordination Number.[a]

		ND/EPSR			FPMD		AnR		HyR	
T [K]		303			all		333	363	333	363
center	shell	r_{max}	r_{min}	N_{coord}	r_{max}	r_{min}	N_{coord}	N_{coord}	N_{coord}	N_{coord}
chloride	choline	4.2	6.7	4.35 ± 1.30	4.2	6.5	3.1 ± 0.6	3.9 ± 0.6	2.9 ± 0.5	3.2 ± 0.6
chloride	urea	4.0	5.5	4.16 ± 2.02	4.1	5.3	3.8 ± 0.8	3.1 ± 0.7	2.6 ± 0.5	2.2 ± 0.5
chloride	water	—	—	—	3.1	3.9	—	—	2.1 ± 0.5	2.2 ± 0.4
choline	urea	5.4	6.9	5.91 ± 2.84	5.1	7.1	8.6 ± 0.7	8.0 ± 0.7	6.8 ± 0.5	5.9 ± 0.7
choline	water	—	—	—	4.9	6.6	—	—	8.3 ± 0.5	8.7 ± 0.7
urea	urea	4.3	6.1	6.77 ± 3.05	4.7	6.3	4.9 ± 0.5	5.7 ± 0.7	4.4 ± 0.5	5.0 ± 0.6
urea	water	—	—	—	3.7	5.7	—	—	5.0 ± 0.3	5.2 ± 0.5
water	water	—	—	—	2.8	3.7	—	—	1.9 ± 0.3	1.3 ± 0.3

[a]For unlike pairs, N_{coord} is provided only for the center molecule being preferentially the H-bond acceptor; the corresponding N_{coord} for the reverse order can be determined from the mole ratio. For the FPMD data, the positions of the maxima and minima reflect averages and the combined uncertainties from block averages and determination of r_{min} are given.

Addition of water leads to a large decrease in the first-peak height and a 30% reduction in N_{coord} for the chloride–urea pair and a smearing out of the RDF for the chloride–chloride pair, i.e., water preferentially binds to the anion. In contrast, the first-peak height for the chloride–choline pair is slightly larger and N_{coord} is reduced by only 10% for HyR compared to AnR. The first-peak height for the water–water RDF in HyR is somewhat higher than for bulk water [193], but $N_{\mathrm{coord}} \approx 1.6$ indicates that this is only due to a dilution effect.

For the discussion of atom–atom RDFs, we focus mostly on those involving a polar H atom as one member of the pair. Each urea molecule contains four H atoms, two of which are in the *trans* orientation with respect to the O atom, while the other two are in the *cis* orientation; these are denoted in the following as $\mathrm{H}_{t\mathrm{U}}$ and $\mathrm{H}_{c\mathrm{U}}$ (see Figure 5.1). It was shown in previous force-field-based (FFB) MD studies [171] that the $\mathrm{H}_{t\mathrm{U}}$ and $\mathrm{H}_{c\mathrm{U}}$ atoms have noticeable differences in their surroundings, and therefore their structural properties must be probed separately. Analysis of the FPMD trajectories shows no occurrences of dihedral rotations around the N–C bonds of urea which allows for straightforward assignment of $\mathrm{H}_{t\mathrm{U}}$ and $\mathrm{H}_{c\mathrm{U}}$ atoms.

Figure 5.3 illustrates the RDFs between the two types of H_{U} atoms and either the O atom of urea (O_{U}) or the chloride ion. The H_{U}–O_{U} RDFs exhibit two distinct peaks at $r = 1.9$ and 3.5 Å that were also observed in the ND/EPSR and FFBMD studies (all comparisons are made here to the 0.8Amb variant; note that the FFBMD RDFs also include intramolecular contributions leading to additional peaks at 2.4 and 3.0 Å for $\mathrm{H}_{c\mathrm{U}}$ and $\mathrm{H}_{t\mathrm{U}}$, respectively). The first peak at $r = 1.9$ Å arises from a direct H-bond and the geometry of the NH_2 group leads to the second peak at 3.5 Å. This second peak is pronounced compared to other H-bonding liquids because two urea molecules often form two H-bonds (either bifurcated involving two *trans*-H atoms or via an eight-membered ring involving two *cis*-H atoms) that constrain the relative orientation of the pair of molecules. The number integrals also indicate that a significant fraction of the O_{U} atoms is surrounded by more than one H_{U} atom in the first shell. The heights of the first peak in the H_{U}–O_{U} RDFs obtained from the FPMD simulations fall in between the ND/EPSR and FFBMD RDFs, and overall the ND/EPSR data shows the least extent of structuring. Furthermore, both FPMD and FFBMD yield a taller first peak for $\mathrm{H}_{c\mathrm{U}}$, whereas there is no significant difference in peak heights for the ND/EPSR

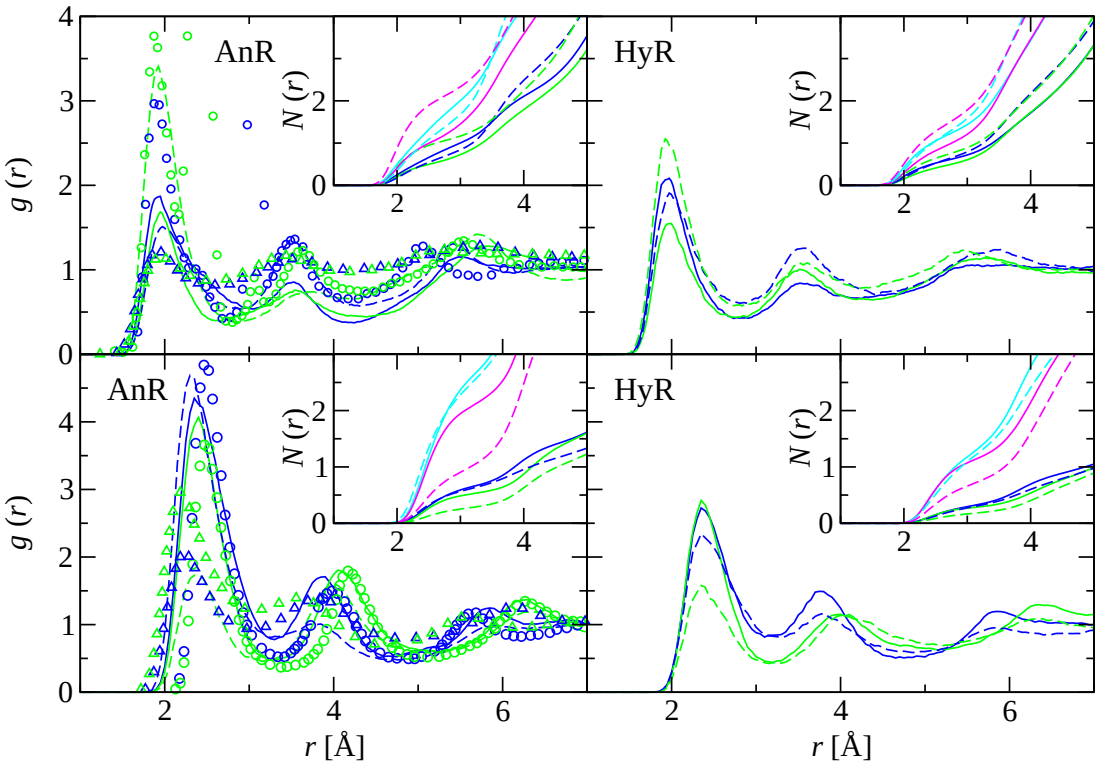

Figure 5.3: H_U–O_U (top) and H_U–Cl (bottom) RDFs and corresponding number integrals (insets) for AnR (left) and HyR (right). The solid and dashed lines show the FPMD simulations at $T = 333$ and 363 K, respectively, and triangles and circles represent RDFs from ND/EPSR ($T = 303$ K) [169] and FFBMD simulations ($T = 298$ K) [171], respectively. Blue and green indicate RDFs and number integrals with H_{tU} and H_{cU} atoms, respectively, as the center, and cyan and magenta lines correspond to number integrals with H_{tU} and H_{cU} atoms, respectively, as the shell (the H_U/O_U and H_U/Cl atom ratios are 2 and 4, respectively).

data. Considering the statistical uncertainties from the short trajectories, the FPMD simulations do not show a significant temperature effect. The H_U–O_U RDFs are also not affected by addition of water, but the corresponding number integrals exhibit a significant downward shift.

In contrast, the differences between the AnR and HyR systems are much more pronounced for the H_U–Cl pairs; i.e., addition of water significantly alters the first solvation shell around chloride ions. Due to the larger size of the chloride ion, the

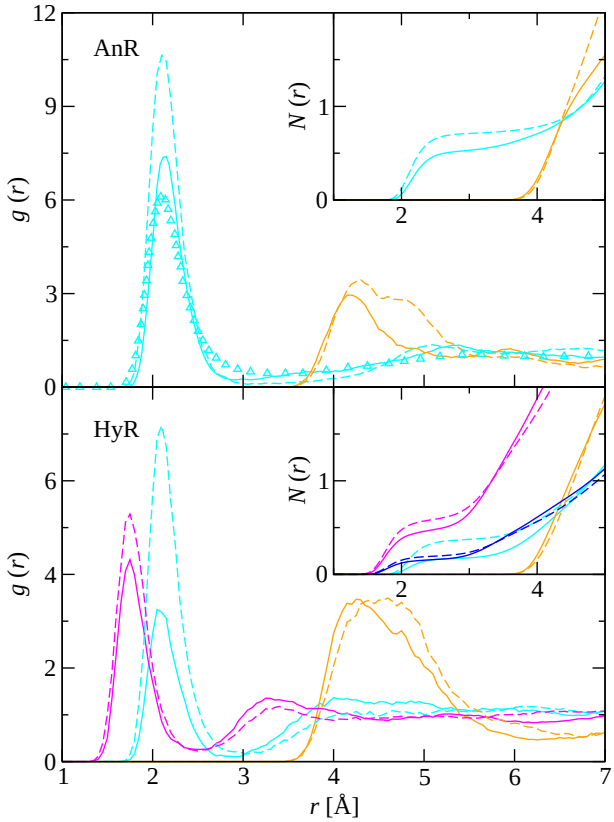

Figure 5.4: H_C–Cl (cyan), N_C–Cl (orange), and H_C–O_W (magenta) RDFs and corresponding number integrals (also O_W–H_C (blue)) for AnR (top) and HyR (bottom). The solid and dashed lines show the FPMD simulations at $T = 333$ and 363 K, respectively, and the triangles denote the ND/EPSR data (303 K) [169].

peak positions (at 2.4 and 2.2 Å for FPMD and ND/EPSR, respectively) are shifted to slightly larger distances compared to those involving O_U as the acceptor, but otherwise the features are remarkably similar. Again, the peak heights obtained from the FPMD simulations fall mostly in between those from ND/EPSR and the FFBMD simulations.

Selected RDFs for the choline cation are shown in Figure 5.4. Only the hydrogen atom belonging to choline's hydroxyl group can act as donor in an H-bond and is denoted here as H_C. Comparing the H–Cl RDFs for urea and the choline cation in AnR, it is notable that the first peak (located at 2.1 Å for FPMD and ND/EPSR) is higher for H_C than for both types of H_U; this feature is also observed in the ND/EPSR data

Table 5.2: Average Fractions of Choline–Chloride Pairs with Contacts through Either H_C, N_C, or Chelating to Both.[a]

system	T [K]	H_C	N_C	both
AnR	333	0.18_3	0.66_4	0.16_4
	363	0.05_2	0.82_3	0.13_4
HyR	333	0.10_4	0.66_3	0.24_5
	363	0.08_3	0.75_4	0.17_4

[a] Subscripts represent the standard deviation in the last digit.

[169]. The value of the plateau in the corresponding number integrals is similar for both types of polar hydrogen atoms, but the plateau is much more pronounced for H_C than for H_U. On average, there is about 0.5 of a Cl anion in the first solvation shell of H_C and vice versa. Nevertheless, considering that the H_C:Cl and H_U:Cl ratios are 1:1 and 8:1 in reline, the solvation shell of the Cl anion is mostly populated by H_U. Addition of water to the DES, as in the case of H_U–Cl, disrupts H_C–Cl contacts (the number integral decreases about twofold) due to formation of H-bonds between water and choline's hydroxyl group or water and chloride. Of interest for the DES structure is also the N_C–Cl RDF. Since the quaternary nitrogen atom is well shielded by three methyl groups and one methylene group, the first peak is found at a relatively large distance of 4.2 Å; however, this structuring is much more pronounced (with a height of about 3) than for any other atom–atom RDF at r between 4 and 5 Å. Interestingly, the N_C–Cl peak height and number integral are not much influenced by addition of water. From the RDFs (see Figure 5.4), one cannot obtain information on whether some choline cations "chelate" a chloride anion through simultaneous interaction with both H_C and N_C. To this extent, distance cutoffs at 3.0 and 6.0 Å for H_C–Cl and N_C–Cl contacts, respectively, are used to identify strong contacts. The numerical data provided in Table 5.2 indicate that a chelate-type interaction is present for only about one sixth of the choline–chloride pairs, whereas about 70% of the pairs form via only an N_C–Cl contact (with H_C–Cl contacts accounting for the remainder). These features illustrate a complex structure where a given choline cation tends to form a strong H-bond via H_C to one chloride ion, but also interacts strongly with the same or other chloride ions

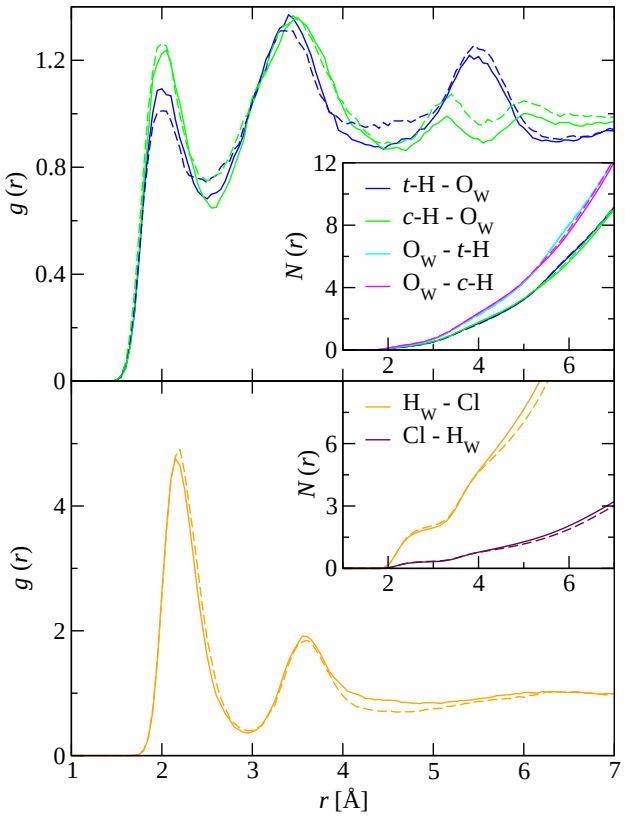

Figure 5.5: H_U–O_W (top) and H_W–Cl (bottom) RDFs and corresponding number integrals. The solid and dashed lines correspond to 333 K and 363 K, respectively. Blue indicates *trans*-H and green *cis*-H.

through the quaternary ammonium group.

Moreover, the addition of water introduces new H-bonding moieties, which can bond not only to the H atoms of a H-bond donor (urea) but also to the chloride ions. Structures of these interactive pairs are given through the corresponding RDFs and number integrals, which are shown in Figure 5.5. It can be seen that H_U–O_W RDFs are very similar for both H_{tU} and H_{cU} suggesting that both hydrogen atoms are equally available for H-bonding to water. Additionally, there is a very defined structure for H_W–Cl pairs showing that, with the addition of water, chloride ions are now also bound to new H-bond donors. Temperature is found to have only a minor effect on the aforementioned RDFs.

5.3.2 Hydrogen-Bond Analysis

A strong and varied H-bond environment between the cation donor, anion acceptor, and neutral H-bond donor molecules is considered pivotal for the suppression of the melting point leading to the formation of DESs. Thus, a detailed analysis of the H-bond environment is provided here. Previously, strict criteria for the determination of an H-bond were used for AnR and similar systems [171, 194]; that is, an H-bonded pair needs to satisfy a distance of smaller than 3.5 Å between donor and acceptor (X and Y) heavy atoms and an X–H$\cdots$Y angle of greater than 150°. However, as originally suggested by Wernet et al. [195] for liquid water, such a set of two separate criteria (one distance and one angle) may not be adequate to capture the presence of a (strong) H-bond. Wernet et al. introduced a coupled H-bond definition, where the distance and the angle are connected to define an ellipse. In order to assess this approach for the AnR and HyR systems (and potentially also other DESs), combined radial-angular distribution functions (RADFs) were computed for all possible pairs of H-bond donors and acceptors (see Figure 5.6). It is clear that the peaks in all RADFs are indeed much better described by an elliptical boundary than a rectangular boundary. Thus, H-bond criteria were defined here using elliptical dependencies for the X–Y distance and the X–H$\cdots$Y angle. Given the similar areas occupied by the peaks in the 12 RADFs, a common half-width of 0.50 Å is used for the X–Y distance and a common half-height of 45° is applied for the X–H$\cdots$Y angle. For the latter, the ellipse is, of course, always centered at 180°, but the center for the X–Y distance is found to depend on the donor and acceptor atom identities. In this work, the following values are used for the X–Y distances of the ellipse centers: 2.90 for Å O–O, 3.05 Å for N–O, 3.25 Å for O–Cl, and 3.42 Å for N–Cl pairs.

From the data presented in Figure 5.6, it is evident that the requirement of the X–H$\cdots$Y angle being greater than 150° is indeed much too strict for all but the O_W–O_C pair (and the O_C–O_C pair that does not form any H-bonds due to the electrostatic repulsion between two cations). The situation is more varied for the distance criterion. Although an upper bound of 3.5 Å is reasonable for the N–O pairs, it is much too strict for N–Cl and too strict for O–Cl, and, at the same time, too permissive for O–O pairs. Clearly, a common H-bond definition with a rectangular bounding box leads to an overestimation of the number of H-bonds for O–O pairs and an underestimation for

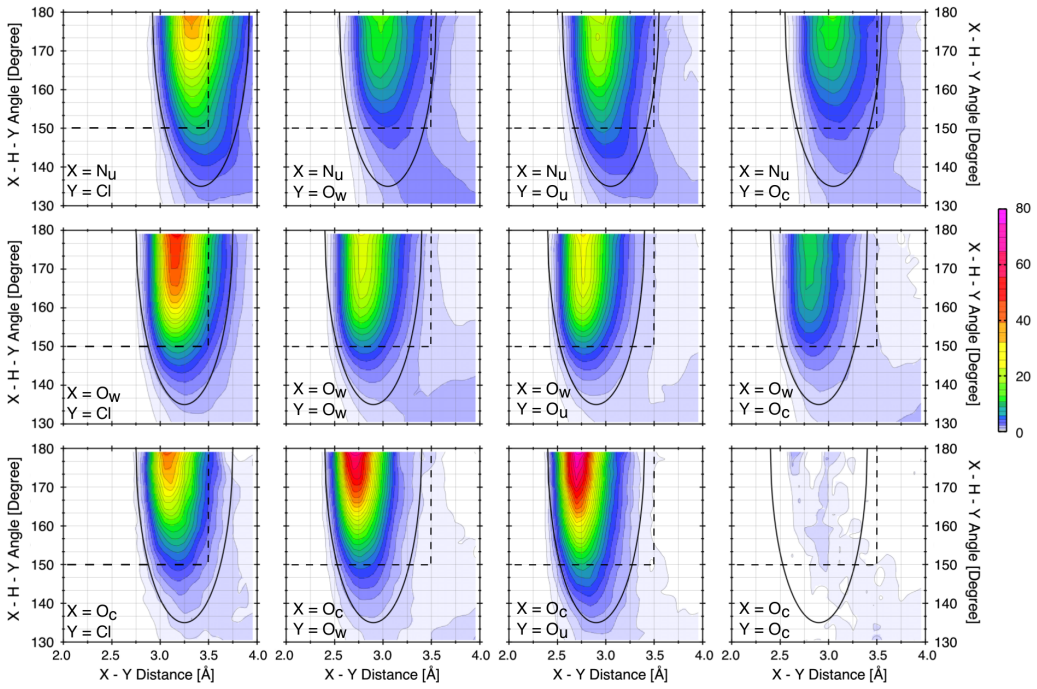

Figure 5.6: RADFs for all combinations of H-bond donor (X) and H-bond acceptor (Y) present in AnR and HyR. The dashed lines show the rectangular boundaries for the H-bond criteria used by Perkins et al. [171], and the solid lines show the elliptical boundaries used to define the different types of H-bonds in this work.

all other pairs. Therefore, use of the Wernet-like H-bond definition with ellipse centers reflecting the specific donor–acceptor combination is strongly recommended for analysis of DESs (and any other system containing acceptor atoms differing in atomic radii).

Using the set of elliptical H-bond criteria, the fractions of H-bonds per H_U atom were determined and are listed in Table 5.3. The average fractions obtained here differ markedly from those computed for AnR from FFBMD simulations [171] (e.g., 0.09 for H_{tU}–Cl and 0.06 H_{cU}–Cl that are about 5 times smaller). Considering the overall good agreement for RDFs, these differences are likely caused by the rectangular criterion being much too strict for these specific donor–acceptor combinations (see Figure 5.6).

For the discussion of the H-bond network, it is important to note that the reline system provides a slight imbalance of H-bond donors and acceptor sites per 1:2 formula unit: 1 H_C and 8 H_U versus 2 on O_C, 4 on O_U, and 4 on the chloride anion. As already

Table 5.3: Average Fractions of H-Bonds of the Type H_U–X.[a]

		H_{tU}					H_{cU}				
system	T [K]	O_U	O_C	Cl	O_W	free	O_U	O_C	Cl	O_W	free
AnR	333	0.23_4	0.01_1	0.44_4	—	0.32	0.22_4	0.12_3	0.40_5	—	0.26
	363	0.21_5	0.01_1	0.44_6	—	0.34	0.34_6	0.03_2	0.26_4	—	0.37
HyR	333	0.23_5	0.04_2	0.22_4	0.20_5	0.31	0.16_5	0.07_4	0.22_5	0.21_5	0.34
	363	0.14_5	0.03_3	0.19_4	0.16_5	0.48	0.17_6	0.05_3	0.12_4	0.17_5	0.49

[a] Subscripts represent the standard deviation for the last digit.

deduced from the corresponding RDFs, H_{tU} atoms show slightly higher propensity for H–Cl H-bonding, while H_{cU} atoms exhibit more H–O_{urea} H-bonding, which agrees with the previous FFBMD studies [171]. Interestingly, this trend is enhanced when the temperature of either system is increased. Moreover, H_{tU} atoms are only very rarely involved in H-bonding to choline O atoms, whereas about 10% of the H-bonds formed by H_{cU} are with O_C as acceptor. The addition of water to reline decreases the fraction of H_{urea}–Cl H-bonds by a factor of about 2, and this can be explained by the anions H-bonding to water molecules (on average, a chloride ion forms 1.7 H-bonds with water). The addition of water leads to only a small decrease by on average a factor of 1.2 for H_{tU}–O_{urea} and H_{cU}–$O_{choline}$ H-bonds, but leads to a larger decrease by a factor of 1.7 for H_{cU}–O_{urea} H-bonds. This is likely caused by the H_{cU} atoms being slightly more shielded than the H_{tU} atoms. Both types of H_U atoms have the same fraction of H-bonds with water, which is consistent with the RDFs being very similar. Overall, the addition of water leads to hydrated components of the DES causing weaker interactions between the species (decreased number of H-bonds), which is consistent with previous FFBMD studies [174].

Increasing the temperature of the system leads to a slight disruption of the H-bond network and the fractions of H_{tU} and H_{cU} atoms not involved in formation of a H-bond increases by factors of 1.3 and 1.4, respectively.

Switching attention to the chloride ion as another structure defining component, all possible H-bonds between Cl and H were analyzed (see Table 5.4). For AnR, the average numbers of H-bonds per chloride ion are 3.9 at 333 K and 3.7 at 363 K. The majority of chloride ions are bonded to either three – five H_U atoms or to two – four H_U atoms and one H_C atom (69% at 333 K and 51% at 363 K). For HyR, the average

Table 5.4: Fraction of Chloride Ions and Urea O Atoms that are H-bonded to a Specific Combination of Species. n_{UCW} Denotes the Number of Bonded Urea Molecules, Choline Cations, and Water Molecules, Respectively.[a]

Acceptor	n_{UCW}	AnR		HyR	
T [K]		333	363	333	363
Cl	102	0.00	0.00	0.11	0.10
	110	0.15	0.12	0.00	0.01
	202	0.00	0.00	0.12	0.09
	210	0.07	0.10	0.01	0.01
	300	0.11	0.17	0.01	0.00
	310	0.11	0.05	0.00	0.00
	400	0.16	0.13	0.01	0.01
	500	0.13	0.07	0.00	0.00
O	000	0.21	0.16	0.08	0.05
	001	0.00	0.00	0.19	0.18
	010	0.12	0.05	0.03	0.00
	100	0.38	0.33	0.18	0.11
	101	0.00	0.00	0.16	0.28
	200	0.20	0.31	0.09	0.13

[a] Data are only included when the mean of both temperatures for a system yields a fraction of 0.08 or larger.

numbers of H-bonds per chloride ion are 3.6 at 333 K and 3.4 at 363 K. In this case, the distribution of various species bonded to the chloride ion is much more diverse (and the environments included in the table account for only about 20%), but the major combinations are one – three H_U atoms and one – two H_W atoms (52% at 333 K and 44% at 363 K). Typical solvation environments for the chloride ions with three or four H-bond donors are illustrated in Figure 5.7. These snapshots also show that some of the H_U atoms are distorted away from the usual planar urea geometry in order to form the H-bonds.

A similar analysis is carried for the O_U atoms (see bottom part of Table 5.4). The distribution of species is very different in this case because O_U possesses only two accessible acceptor sites compared to four sites for the chloride ion and O_U is also a much weaker H-bond acceptor. Particularly in AnR, a considerable fraction of O_U is not bound to any H-donors. In AnR, formation of a single H-bond to another urea

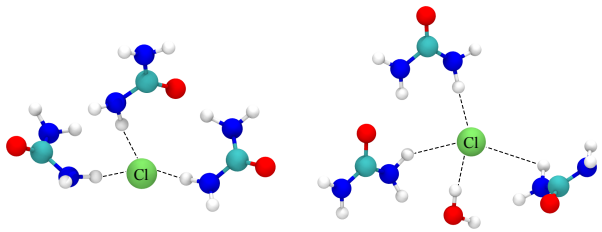

Figure 5.7: Examples of the solvation environment of the chloride ion for AnR (left) and HyR (right) at $T = 333$ K. Cl is shown in lime, C in cyan, N in blue, O in red, and H atoms in white.

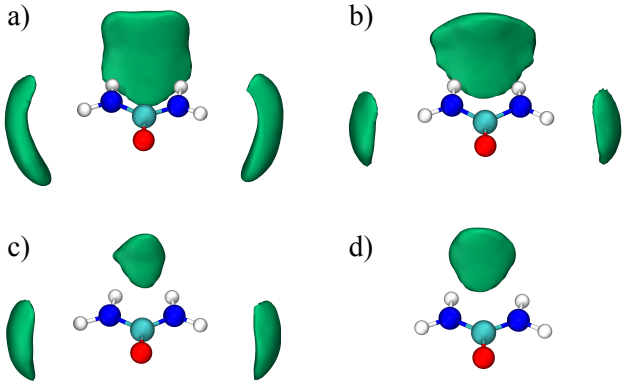

Figure 5.8: Spatial distribution of chloride ions around urea molecules: (a) AnR at 333 K, (b) AnR at 363K, (c) HyR at 333 K, and (d) HyR at 363 K. The isosurfaces encompass the regions with a density at least twice as high as the uniform density.

molecule is most prevalent, followed by two H-bonds to urea or one H-bond to choline. In HyR, H-bond formation to either one urea and one water molecule or to only one water molecule occur each in about 20% of the local environments.

In addition to the quantitative analysis, spatial distribution functions of chloride ions around urea molecules are computed using the *TRAVIS* [196] package. This analysis illuminates where the anions are preferentially localized and the extent to which the close-range structure is influenced by temperature and the addition of water as a co-solvent (see Figure 5.8). From the depicted isosurfaces, it is clear that increasing the temperature significantly decreases the density of chloride ions around the H_{cU} atoms, but affects the density of chloride ions around the H_{tU} atoms to a lesser extent. Moreover, the addition of water depletes anion density around both types of H_U atoms.

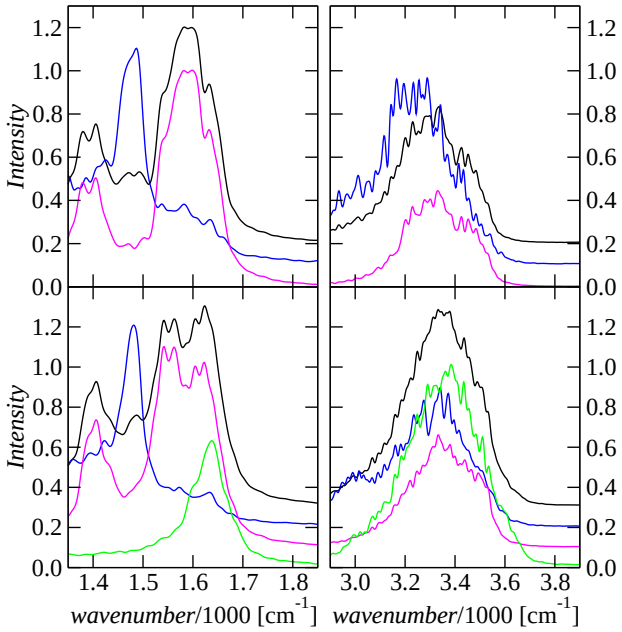

Figure 5.9: Computed infrared spectra for the DES system (top) and its mixture with water (bottom). Black represents the combined spectrum, magenta the contribution from urea, blue the contribution from the choline cation, and green the contribution from water. Some lines are offset vertically for clarity.

5.3.3 Infrared Spectra

Infrared (IR) spectroscopy is capable of providing information about extent and strength of H-bonding. Typically, vibrational frequencies can be obtained from MD simulations using conventional Hessian diagonalization [197], but this approach cannot be accurately applied to bulk systems. Thus vibrational frequencies were obtained from maximally localized Wannier centers [198], which were computed every five time steps using the Fourier transform of the dipole autocorrelation function in the *TRAVIS* package [199]. Due to the intrinsic *ab initio* nature of the simulations, this approach directly accounts for anharmonicity.

As noted in the experimental paper [171], the most interesting regions of the spectrum are 1600–1700 cm^{-1} and 3000–3600 cm^{-1}. The first region corresponds to NH_2 bending and C=O stretching modes, while the second region corresponds to N-H and

O-H stretching modes. Computed IR spectra for both systems, in addition to contributions of the individual components, are given in Figure 5.9. It is apparent that N-H and O-H stretching bands are too broad to distinguish specific features in the main peak, but it is clear that the addition of water leads to a shift to higher frequencies, which can be explained by stronger H-bonds involving water instead of urea as the acceptor. The peak positions corresponding to the NH_2 bending and C=O stretching modes of urea agree well with the experimental measurements and are the main contributors to the overall DES spectrum in this region. The addition of water leads to peak splitting for the NH_2 bending mode suggesting that the amine groups now experience a more heterogeneous environment.

5.3.4 Transport Properties

DESs (and also ILs) are highly viscous and, hence, studying their transport properties using MD simulations requires very long trajectories or special simulation techniques to accelerate convergence [200, 171]. Given the relatively short times currently accessible for FPMD simulations, all quantitative predictions should be interpreted with caution [201, 181]. Prior to calculation of self-diffusion coefficients, it is important to assess whether the different species are found in the diffusive regime for a sufficiently long time interval, i.e., the slope in a log–log plot of the mean square displacement (MSD) versus time is equal to unity.

Calculated MSDs (averaged over multiple time origins during the production period) [4] are presented in Figure 5.10. Rapid changes in the slope for $t > 80$ ps are due to insufficient statistics. For AnR at 333 K, none of the MSDs for the three species reaches 1 $Å^2$, i.e., the particles are mostly rattling in their solvation cage. In comparison, the MSD values for the chloride ion, urea, and choline are about one order of magnitude larger when either the temperature is raised to 363 K for AnR or upon addition of water (HyR system) at 333 K. However, with the exception of water, the MSDs for none of the species in AnR at 363 K and HyR at 333 K exhibit a slope of unity in the log-log plots for the $\approx$ 100 ps time frame accessed in the current FPMD simulations, which is not surprising because the MSDs do not extend to values larger than the square of the COM peak positions (see Figure 5.2). For HyR, increasing the temperature from 333 to 363 K, nearly doubles the MSD values reached at 80 ps for

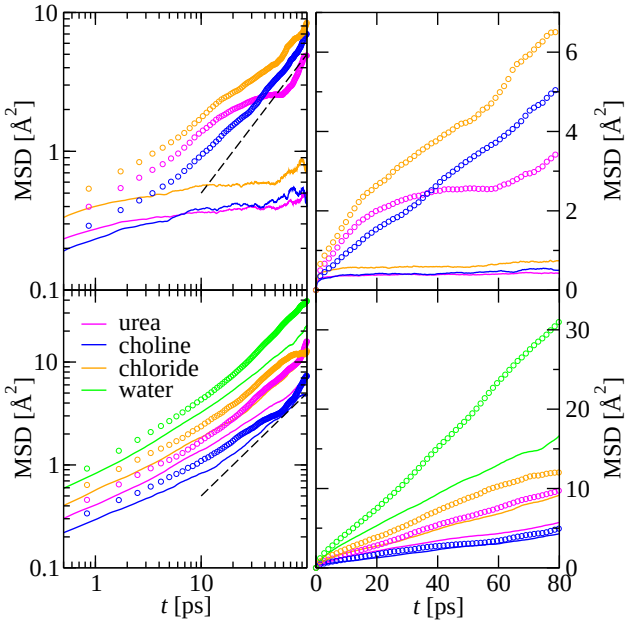

Figure 5.10: Mean square displacements for all species in AnR (top) and HyR (bottom). Lines and circles represent data at 333 and 363 K, respectively. The black dashed lines in the log-log plots indicate a slope of unity.

water, chloride ion, and urea, but not for choline. At the higher temperature, the MSD slope in the log–log plot for water reaches unity at $t \approx 20$ ps, with the slopes for the chloride ion and urea also being very close to unity. Water is clearly the most mobile species in HyR, and the chloride ion appears to be somewhat more mobile than the larger urea and choline species, but the MSDs for chloride and urea cross at $t \approx 90$ ps for HyR at 363 K. Diffusion coefficients for HyR at 363 K can be cautiously estimated using the Einstein relation [4] across the time interval from 20 to 80 ps, and the values are 7×10^{-10} m^2/s for water and 2×10^{-10} m^2/s for both chloride ion and urea. In comparison, the experimental values [202] for anhydrous reline at 328 K are 2.10×10^{-11} and 3.55×10^{-11} m^2/s for choline and urea, respectively, i.e., supporting that hydration significantly increases mobility of all species. Values of 1.1×10^{-10}, 0.97×10^{-10}, and 0.87×10^{-10} for water, urea, and chloride ion were observed in FFBMD simulations [174] of a 10 wt% reline–water mixture at 298 K (compared to 17 wt% mixture studied in this work); i.e., confirming the sluggish behavior of HyR at lower temperature.

5.4 Conclusion

Despite the relatively short accessible time scales, FPMD simulations can provide significant insights on structure, speciation, H-bond aggregation, and vibrational and, with caution, transport properties of DESs. The current FPMD simulations for AnR and HyR show that these liquids are microheterogeneous with formation of preferred H-bonded aggregates. In AnR, the most significant structure determining factor are the H-bonds formed between urea and chloride ions. The addition of water disrupts the H-bonding between urea and the anion by preferentially solvating the anion and replacing some of the urea–chloride H-bonds with urea–water H-bonds. This disruption of the urea–chloride aggregation leads to a significant increase of mobility for these two species. Overall, the structural features of the anhydrous reline agree well with experimental data [169] and MD simulations using molecular mechanics force fields [171]. However, we also show the importance of accurate bonding criteria for the analysis of H-bonds in complex liquid systems, especially, when they contain acceptor atoms differing in atomic radii.

Chapter 6

Mercury Capture from Petroleum Using Deep Eutectic Solvents

6.1 Introduction[1]

Energy efficient, economically feasible and environmentally friendly processes for the production of fuels and chemical feedstocks are highly desirable. In particular, carbon dioxide capture and natural gas sweetening technologies have garnered significant research efforts over the past two decades [80, 203, 204]. On the other hand, comparatively little effort has been directed towards mercury (Hg) capture due to its very small concentration that varies between 0.01 ppb and 10 ppm, depending on the geological location of the oil/gas reservoir [205]. Mercury in crude oil/natural gas is present in different toxic species: elemental mercury (Hg^0) is most prevalent, but mercuric halides (mostly, $HgCl_2$), organic mercury compounds (RHgR and RHgCl) and mercury-sulfur complexes can also be found [205]. Beyond mercurys health and safety risks for the biosphere, mercury is also a major problem in oil and gas processing units as it deposits in the cryogenic units forming amalgams with different metals (e.g., aluminum) that lead to equipment degradation, and it poisons many catalysts [205]. Additionally, Hg

[1]This chapter builds upon the previous chapter and represents some outcomes of a research project, where experiments were carried out by S. E. E. Warrag, D. J. G. P. van Osch, M. C. Kroon, and C. J. Peters and simulations were carried out by E. O. Fetisov in collaboration with D. B. Harwood and advised by J. I. Siepmann. A report on this project has been submitted to a journal. S. E. E. Warrag, M. C. Kroon, and C. J. Peters are co-inventors on a patent application relating to this work.

emissions are a major environmental concern and are classified as hazardous air pollutants [206]. Owing to mercurys adverse environmental effects, as well as the operational issues in the oil and gas processing industry, it is rendered mandatory to develop an efficient removal process.

Several technologies are commercially available for mercury capture from liquid/gaseous hydrocarbon streams based on either amalgamation [207], physical adsorption [208], chemical adsorption or reactive absorption [209]. The most mature technologies are adsorption on activated carbon and on sulfur/transition metal sulfides impregnated on a solid support, such as activated carbon, alumina, zeolite or silica [209, 210, 211]. Due to the sensitivity of sulfur to moisture in organic systems, the latter is less suitable for application in liquid streams [210]. Some other technologies employ regenerative molecular sieves impregnated with silver, but this is an expensive option compared to activated carbon beds [210]. Recently, Clariant and Petronas announced the commercialization of a new solid-supported ionic liquid (SSIL) mercury removal technology [212]. SSIL effectively removes elemental, organic and inorganic mercury from natural gas [213]. Other ILs also show promise as mercury capture technology [213, 214, 215, 216], but ILs are a relatively expensive class of solvents, and economical regeneration approaches have not yet been reported.

In 2003, Abbott et al. [158] introduced an innovative class of solvents, the so-called deep eutectic solvents (DESs), which could be potential alternatives to ILs. DESs consist of at least one compound acting predominantly as hydrogen-bond donor (HBD) and another compound acting predominantly as hydrogen-bond acceptor (HBA) that form a mixture exhibiting a significantly lower freezing point than both of the pure compounds. DESs have similar properties as ILs (including low vapor pressure, wide liquid range, and low flammability), but are generally much easier to prepare from low-cost HBD and HBA ingredients that that are mostly biodegradable [158, 159, 160, 217, 218].

DESs have already been utilized in a number of separations relevant to oil and gas industries, such as desulfurization, dearomatization, and sweetening [76], and also for gas separation [219] and water reclamation [220]. Additionally, the application of DES-functionalized carbon nanotube adsorbents for mercury removal has been studied [221]. In this work, the first use of DESs as extracting agents for the removal of mercury from Hg-containing liquid hydrocarbons is reported. First principles molecular dynamic

simulations were used to obtain molecular-level information on the mercury solvation.

6.2 Summary of experimental findings

The selection of an extraction solvent for mercury from hydrocarbon liquids depends on four factors: (i) its strong affinity for solvating various mercury species, (ii) its low mutual solubility with hydrocarbons, (iii) its thermal stability, and (iv) its regenerability. Here, the first three factors are addressed. DESs are very polar and, hence, their mutual solubility with aliphatic hydrocarbons is expected to be very low [222]. *n*-Dodecane was selected to represent the aliphatic hydrocarbons in petroleum. It is known that halogen- (particularly, Cl) and nitrogen-containing ILs exhibit excellent extraction efficiency for mercury [216]. Mancini et al. [223] studied the mechanism of the removal of mercury ions from aqueous solutions using hydrophobic/Cl-containing ILs in the absence of chelating agent and suggested that Hg^{2+} ions are transferred to the IL phase through the formation of polyanion species $HgCl_n^{?n+2}$ (where n is 1 to 4) and then extraction by the IL. Cheng et al. [224] evaluated the extraction of Hg^0 from flue gas using 1-alkyl-3-methylimidazolium chloride ILs and identified the formation of a complex [IL cation]$HgCl_3$ on a solid adsorbent by means of Raman and UV-Vis spectroscopies. Thus, salt based/polar DESs are expected to be good candidates for this application. Therefore, the selected DESs were (i) choline chloride:urea (DES-1), (ii) choline chloride:ethylene glycol (DES-2), (iii) choline chloride:levulinic acid (DES-3), and (iv) betaine:levulinic acid (DES-4), where the molar ratio is 1:2 in all four cases. DES-4 was chosen to test the influence of replacing a salt-based HBA with a zwitter-ionic HBA. The extraction efficiency for the system [*n*-dodecane + Hg^0 + DES] was determined by direct solvent-feed extraction in ratios of 1:1 and 2:1 at $T = 303.15$ and 333.15 K and atmospheric pressure. These experimental operating conditions were chosen based on the conventional processing temperatures for the Hg adsorbent beds [210].

No color change was noticed when the DESs were mixed with the *n*-dodecane solution and the mercury was transferred from the non-polar alkane phase to the polar DES phase. The initial and final mercury concentrations in the *n*-dodecane solution, C_i and C_f, respectively, were measured in triplicate for each sample, and each extraction

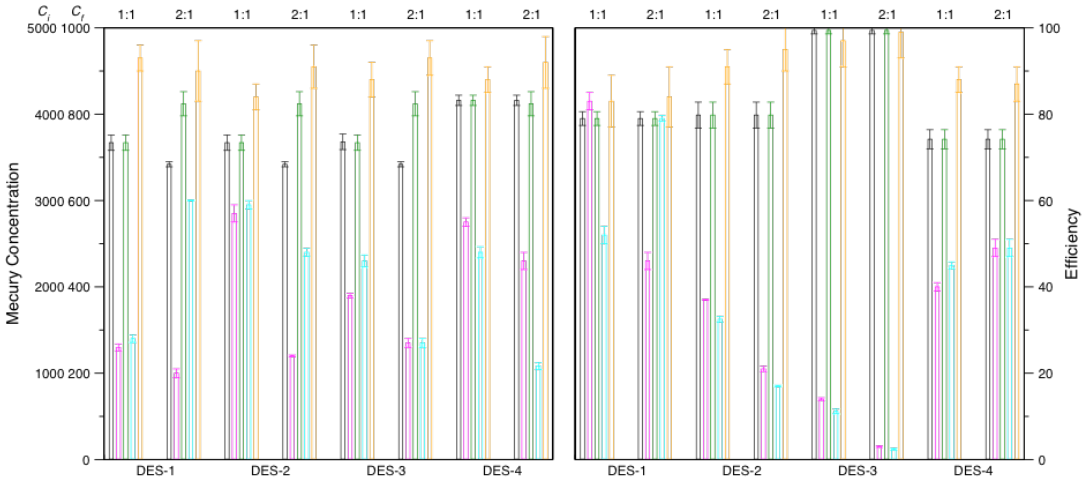

Figure 6.1: Extraction performance data for the four DESs at 303.15 K (left) and 333.15 K (right). The C_i values for the two extraction experiments performed for each solvent and feed ratio are shown as black and green bars using the outer tick labels on the left side. The corresponding C_f values are shown as magenta and cyan bars using the inner tick labels on the left side. The E values averaged over the two extraction experiments are shown as orange bars using the tick labels on the right side.

experiment was done in duplicate. The extraction efficiencies, E, were calculated as follows:

$$E = (C_i - C_f)/C_i \tag{6.1}$$

The results are displayed in Figure 6.1. It should be noted that the statistical uncertainties are fairly large, as one should expect for extraction experiments using very dilute solute concentrations. At 303.15 K, the two extraction experiments yield consistent C_f values when the same feed solution was used. At 333.15 K, however, the C_f values for DES-1 show significant scatter despite use of the same feed solution. Increasing the solvent-to-feed ratio is found to increase the extraction efficiency only for six of the eight cases, but even in those cases the increase in E falls within the statistical uncertainties. All four solvents yield extraction efficiencies in excess of 80% for both temperatures and both solvent:feed ratios. Overall, 9 out of 16 E values fall into the range from 87 to 93%, and E values tend to be somewhat higher at T = 333.15 K with 3 out of 8 satisfying $E \geq 95\%$. Particularly, DES-3 extracts more than 97% of the mercury at 333.15 K, but here a feed solution with an unusually high C_i value was used. The

good extraction performance for DES-1, DES-2, and DES-3 can likely be attributed to the coordination of mercury by Cl anions, but coordination by electronegative oxygen atoms of the solvent may also play a role. For the zwitter-ionic DES-4, coordination by the carboxylate group may provide a favorable environment for mercury. An affinity of Hg atoms to solvate in acidic mediums has been described previously [225]. A molecular view of the solvation in DES-1 is provided in the next section. Gibbs free energies of transfer were also estimated from the concentration data and are found to range from -4 to -14 kJ mol^{-1} (with uncertainties being less than 2 kJ mol^{-1}); thus, solvation in the DESs is clearly much preferred over solvation in a petroleum oil, but the uncertainties preclude a detailed ranking of the DESs investigated here.

6.3 Computational methods and results

First principles molecular dynamics (FPMD) simulations were used to provide molecular-level information on the mercury solvation. Due to the significant expense of FPMD, only DES-1 was investigated at an elevated temperature of 363.15 K to improve the sampling. Three different systems containing 8 choline chloride formula units and 16 urea molecules as the solvent and either 2 Hg atoms, 2 HgCl formula units, or 2 $HgCl_2$ formula units were studied to probe the effects of Hg speciation. The systems were initialized using Monte Carlo simulations with the modified AMBER [171] force field for the solvent and the Universal force field for mercury [226] to obtain initial configurations for the FPMD simulations. Two initial configurations were generated for each system: "short" and "long" with the initial Hg–Hg distances being 2.6 and 5.0 Å, respectively. The Monte Carlo simulations consisted of approximately 3.6 million Monte Carlo moves using translations of the atomic species and translations and rotations for rigid molecular species (choline and urea). The simulations were carried out in the canonical ensemble with the box volumes determined from the molecular volume of DES-1 [189] and the van der Waals or ionic volumes of the mercury species. The resulting box volumes for the Hg, HgCl, and $HgCl_2$ systems were 3018 Å^3, 3047 Å^3, and 3094 Å^3, respectively.

All FPMD simulations were performed with the CP2K simulation package [47],

which solves the Kohn-Sham formulation of density functional theory with the Gaussian plane wave method [48]. The BLYP functional [49, 50] with the third-generation dispersion correction (D3) of Grimme [51] was used along with a triple-zeta, double polarization basis set [52] for non-metal atoms, double-zeta MOLOPT basis set for mercury atoms [52], and GTH pseudopotentials [147, 227].The plane wave cut-off was set to 400 Ry. A time step of 0.5 fs was used, and the temperature was controlled using massive Nosé–Hoover chain thermostats [192, 146]. All systems were equilibrated for 40 ps in the canonical ensemble at the box volumes mentioned above. Thereafter, at least 60 ps of production trajectories were generated for subsequent analysis. Configurations for structural analysis were collected during production runs every ten time steps. Additional simulations in the canonical ensemble with umbrella sampling of Hg–Hg distances were performed for the three systems to calculate the potential of mean force (PMF). The Hg–Hg distances ranged from 2.6 to 7.1 Åwith windows equally spaced by 0.3 Å. A harmonic potential with a force constant of 400 kJ mol^{-1} $Å^{-2}$ was used to constrain the Hg–Hg distance. In each window, the system was equilibrated for 5 ps, and statistics for the PMF were collected over at least 30 ps. To extract the Helmholtz free energy, the weighted histogram analysis method was used [228].

Out of necessity, the FPMD simulations contain Hg in a very high concentration (about four orders of magnitude higher than the experimental systems investigated here). Thus, the first question that needs to be addressed is whether the presence of Hg in these concentrations significantly perturbs the structure of DES-1. Previously, we have already investigated the structure and dynamics of anhydrous and hydrous DES-1 using FPMD simulations (see Ref. [151] and pervious chapter). Figure 6.2 shows the radial distribution functions (RDFs) and the corresponding number integrals (NIs) involving the *trans*- and *cis*-hydrogen atoms of urea as HBD and either Cl anion or the oxygen atom of urea as the HBA. In agreement with previous force-field based simulation studies [171], the FPMD simulations demonstrate a preferential solvation of the Cl anion by the *trans*-hydrogen atoms of urea and double peaks for urea–urea pairs that indicate that the carbonyl oxygen either is involved in two hydrogen bonds with cis-hydrogen atoms from two different urea molecules or in two hydrogen bonds with the two *trans*-hydrogen atoms of the same urea molecule. The small differences between the two systems observed for these RDFs and the corresponding NIs demonstrate that

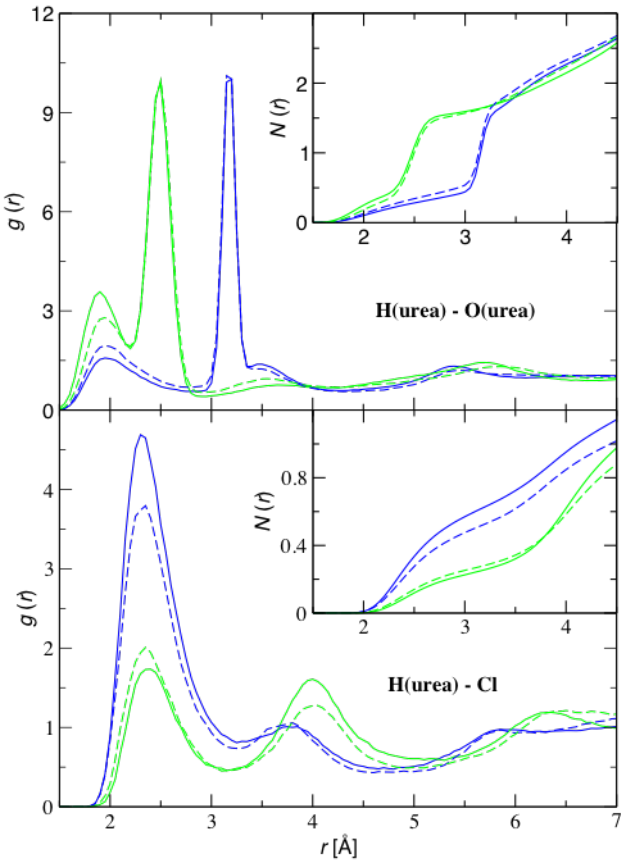

Figure 6.2: Effect of mercury on the structure of DES-1: radial distribution functions and corresponding number integrals for H(urea)–O(urea) and H(urea)–Cl are shown in the top and bottom parts. The solid and dashed lines correspond to systems without and with two additional Hg atoms, respectively. The blue and green lines represent RDFs involving the *trans*- and *cis*-hydrogen atoms of urea.

the solvent structure is not significantly perturbed by the presence of Hg in high concentrations. This may be an indication that the loading capacity in DES-1 may actually be quite high.

Mercury–chlorine RDFs are shown in Figure 6.3 for the systems including pairs of Hg atoms added in three different formal oxidation states and two different initial Hg–Hg distances. For the Hg(0) system, the simulations for the two different initial Hg–Hg distances both yield a peak at $r \approx 3.7$ Å for one Hg atom that is coordinated on average by 1.5 to 2 chlorine atoms as indicated by the NI (see Figure 6.3), whereas the other Hg

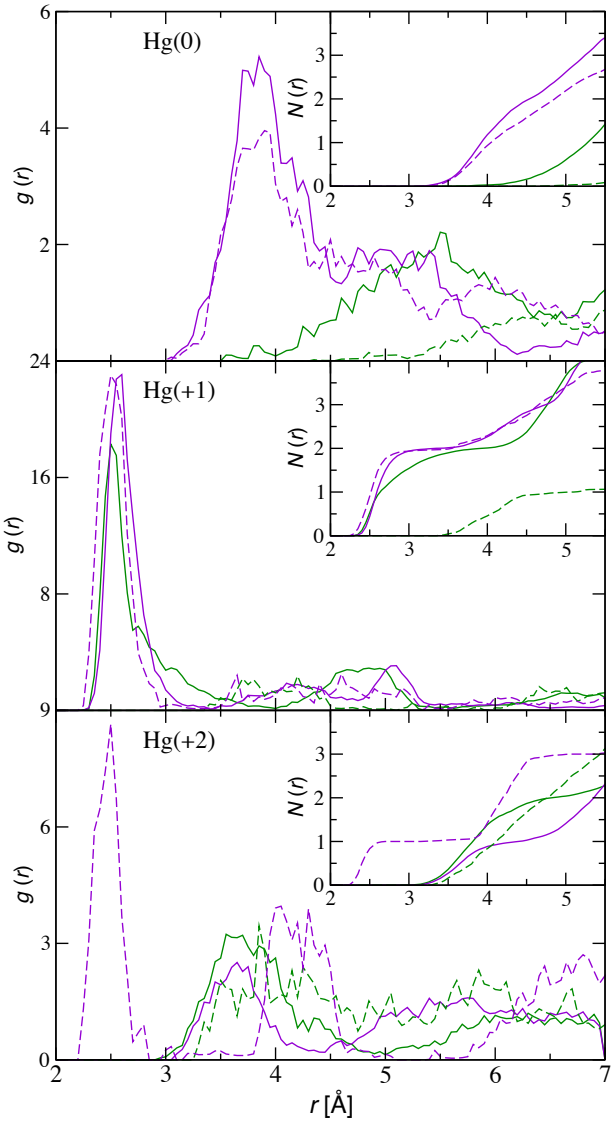

Figure 6.3: Hg–Cl radial distribution functions and the corresponding number integrals for DES-1 systems with the addition of 2 Hg (top), 2 HgCl (middle), and 2 Hg_2Cl_2 compounds (bottom). The solid and dashed lines indicate RDFs for simulations started with the Hg atoms placed either in close proximity or far apart. The purple and green colors distinguish the Hg atoms with the instantaneously tighter and looser coordination by Cl anions.

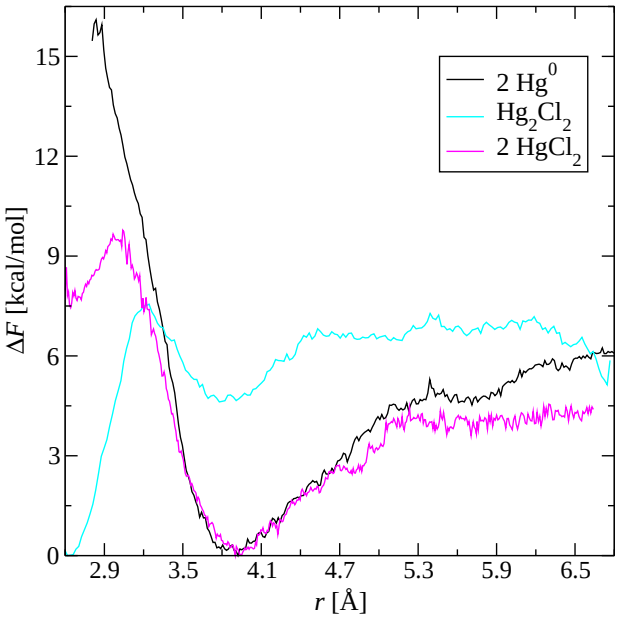

Figure 6.4: Hg–Hg potentials of mean force for DES-1 containing 2 Hg^0, Hg_2Cl_2, and 2 $HgCl_2$

atom is only weakly coordinated by chlorine atoms. In contrast, for the Hg(+1) system, a very sharp peak at $r \approx 2.6$ Å is the most prominent feature that is followed by deep and broad minimum. The NI for this peak is 2.0 and there is a broad plateau due to the minimum in the RDF. This structural feature clearly corresponds to a strongly bound $HgCl_2$ species and is present for both Hg atoms in the simulation started with these atoms in close proximity, but only for one of the Hg atoms in the simulation started with the two Hg atoms at a large distance. Thus, there may be a free energy barrier preventing the two simulations to converge to the same answer. For the Hg(+2) system, both the sharp peak at $r \approx 2.6$ Åand the broader peak at $r \approx 3.7$ Å are present.

To overcome the sampling bottlenecks and to provide more detailed insight on the speciation, potentials of mean force (PMFs) were calculated as function of Hg–Hg separation (see Figure 6.4) for the three systems containing different numbers of Cl atoms that reflect different oxidation states of the mercury species that are extracted from liquid petroleum or natural gas. These PMFs show a global minimum at $r \approx 3.9$ Å for

the Hg(0) and Hg(+2) systems and lend support to the similarity of the Hg–Cl solvation environments for these two systems observed in the RDFs. For the Hg(+1) system, only a weak local minimum is present at $r \approx 3.8$ Å. For the Hg(+2) system, a weak minimum is also found at $r \approx 2.6$ Å. The Hg–Hg species at $r \approx 2.6$ Å correspond to the well-known mercury polycation, whereas the minimum at $r \approx 3.9$ Å corresponds to a pair of Hg species separated by their van der Waals diameter [229]. To characterize the speciation present at these minima, the Hg_2Cl_n aggregation was determined using configurations from the two umbrella sampling windows from 3.5 to 4.1 Å for the Hg(0) and Hg(+2) systems, and using only the window from 2.6 to 2.9 Å for the Hg(+1) system. For the Hg(0) system, Hg_2Cl_3 is the most prevalent speciation with about 40% of all aggregates. Among these Hg_2Cl_3 aggregates, about two thirds have a structure with two Cl atoms coordinating to only one of the Hg atoms, and the other Cl atom either coordinates to both Hg atoms or to the second Hg atom. For the Hg(+2) system, Hg_2Cl_3 is also the most prevalent speciation with about 60% of all aggregates. Here, more than half of these aggregates have a structure with one shared Cl atom and both of the other two Cl atoms coordinating to one of the Hg atoms. Thus, it appears that one of the Hg atoms in the Hg(+2) system is reduced to a lower oxidation state. This conjecture is also supported by the observation that the system contains one Cl_2 species where the two nuclei are separated by less than 2.1 Å, i.e., a distance that would be extremely unlikely for two Cl anions. That is, solvation of Hg(+2) compounds in DES-1 can involve a redox reaction. Based on the uneven coordination of the Hg contact pair, the redox reaction appears to be $2\ Hg^{2+} + 2\ Cl^- \rightarrow Hg^{2+} + Hg^0 + Cl_2$. For the Hg(+1) system, about 40% of the polycation aggregates are found with the $Cl_2HgHgCl_2$ structure, 30% with the $Cl_2HgClHgCl$ structure (where the Cl in the middle coordinates to both Hg atoms), 10% with the $ClHgCl_2HgCl$ structure, and 20% as Hg_2Cl_5 aggregates with various structures.

In summary, the first application of DESs for the removal of elemental mercury (Hg^0) from liquid petroleum with extracting efficiencies exceeding 80%. First principles molecular dynamics simulations indicate that the coordination sphere of mercury is populated predominantly by chloride anions with remarkably similar solvation structures for Hg atom pairs with initially different formal oxidation states.

Chapter 7

Structure and Phase Behavior of Mixed Self-Assembled Alkanethiolate Monolayers on Gold Nanoparticles: A Monte Carlo Study

7.1 Introduction[1]

For many years, various thiolated ligands have been used to tune the properties of gold substrates and nanoparticles. Formation of self-assembled monolayers (SAMs) on the surface of such nanoparticles allows them to be used in various areas ranging from electronics [231] and spectroscopy [232] to biocatalysis [233] and molecular recognition [234]. The most widely used ligand compounds are alkyl thiols and, for most applications, SAMs are formed from mixtures of these compounds [235]. Although SAM-covered nanoparticles have been studied extensively and various preparation methods have been suggested [236, 237], open questions remain pertaining to the structure and

[1]This chapter describes the outcome of a research project carried out by E. O. Fetisov and advised by J. I. Siepmann. A report on this project has been published [230]. Copyright 2016 American Chemical Society.

phase behavior of such systems.

In a pioneering work, Jackson et al. [238] discovered that, for binary mixtures, the alkyl thiol ligands arrange themselves on a gold nanoparticle into stripes that are visible by scanning tunneling microscopy (STM). However, other researchers [239, 240] suggested that the appearance of stripes in the STM images may be attributed to an inadequate analysis approach and instrumental artifacts. Several theoretical studies have been performed on similar systems, including dissipative particle dynamics (DPD) simulations [241] and molecular dynamics (MD) simulations [242] that reveal striped phases for various combinations of two ligand types with different lengths and/or functional groups at the chain terminus. However, these DPD and MD simulations used force field parameters that would lead to bulk liquid–liquid phase separation for alkanes of different lengths that is not observed in experimental studies of liquid alkanes falling into the molecular weight range relevant to SAMs. The DPD-type models with purely repulsive interactions are also not capable of yielding vapor–liquid phase coexistence and those systems would correspond to alkanes in a supernaturally good solvent with solvent–alkane interactions being more favorable than alkane–alkane interactions. Recently, MD simulations with explicit solvent molecules were used to investigate the interfacial behavior of mixed-ligand covered gold nanoparticles in water [243, 244] and aqueous salt solution [245]. Using three different initial spatial distributions of ligands (random, striped, and Janus-like), however, resulted in very different "equilibrated" structures for the coated nanoparticles, i.e., a clear indication that these structures are kinetically trapped and the systems have not reached thermal equilibrium over the length scales accessible to MD simulations. This may not come as a surprise because Bain et al. [246, 247] found that equilibration of mixed monolayers on planar gold substrates can take weeks. It is possible that vertices and edges present on the surface of nanoparticles further hinder the approach to thermodynamic equilibrium.

The motivation for this work is to establish thermodynamic equilibrium structures and to determine the phase behavior for mixed SAMs consisting of alkyl thiol ligands with different degree of length variation coated on gold nanoparticles. To this extent, we use specialized Monte Carlo algorithms that can overcome kinetic barriers and were previously applied to mixed SAMs on flat gold substrates [248]. The model systems investigated here do not include explicit solvent molecules and, hence, correspond to

SAM coated nanoparticles in poor solvent [249].

7.2 Simulation Details

Monte Carlo (MC) simulations for three equimolar alkanethiol mixtures, namely $S(CH_2)_5CH_3/S(CH_2)_9CH_3$, $S(CH_2)_5CH_3/S(CH_2)_{13}CH_3$, and $S(CH_2)_9CH_3/S(CH_2)_{13}CH_3$ (shortened in the following as C6/C10, C6/C14, and C10/C14, respectively) on isolated gold nanoparticles are performed in the canonical (NVT) ensemble. Spherical nanoparticles with three different diameters (50, 70, and 90 Å) are constructed by randomly placing gold atoms on a sphere surface at a density of 10.4 $Å^2$/atom (N_{Au} = 755, 1480, and 2447, respectively, for the three different diameters) and, subsequently, equilibrating the substrate using only translational displacements confined to the sphere surface (here confined translations are achieved by "rotation" of gold atoms using the center of the sphere as origin). The Au–Au potential is taken from Ghorai and Glotzer [242] and the substrate equilibration consists of 200 000 N_{Au} attempted moves at T = 298 K. For the remainder of the simulations, the gold atoms are considered fixed in their positions. Such a model for the nanoparticle was also used by Glotzer and co-workers [241, 242]; filling the nanoparticle with gold atoms would increase the strength of the ligand–substrate interactions and may lead to slightly larger tilt angles.

As the next step, alkanethiolates are placed with their sulfur headgroups in either a random or in a Janus-like arrangement at a distance of 2.38 Å from the gold substrate and with their alkyl tail pointing radially away from the nanoparticle center. The surface coverage is set to 21.6 $Å^2$ per ligand, i.e., the same coverage as found for planar substrates [241, 242, 246] (the numbers of alkanethiolates are N_{lig} = 436, 812, and 1306, respectively, for the three different nanoparticle diameters). Since the densities of gold and sulfur atoms are both specified in Å per atom and the radius of the sulfur sphere is larger than that of the gold sphere, the gold-to-sulfur ratio increases slightly from the 50 Å to the 90 Å nanoparticle (from 1.73 to 1.87, respectively); to some extent, this reflects the greater accessibility of gold atoms on the nanoparticles compared to planar substrates. The resulting coated nanoparticles are equilibrated for at least 100,000 MC cycles (where a cycle consists of N_{lig} randomly selected moves) using only translational, rotational (with the sulfur atom as origin), and configurational-bias Monte

Carlo (CBMC) moves [250, 251] on the alkane thiol chains. Thereafter, CBMC identity exchange moves [248] between chains of different lengths are added to the mix of moves, and each system is equilibrated for another 200 000 MC cycles. For an identity exchange move, two chains of different types are randomly selected and the CH_3 unit of the shorter chain is replaced by a $(CH_2)_mCH_3$ group; concomitantly, the $(CH_2)_mCH_3$ group of the longer chain is replaced by a CH_3 unit, where m is the difference in chain length. During this move, the conformation of the lower part of the ligands is not altered, but the groups being exchanged are generated by the CBMC scheme. The CBMC identity exchange move efficiently samples the spatial distribution of ligands by overcoming the large kinetic barriers that inherently hamper the approach to equilibrium via diffusion on the substrate or hopping events (desorption, diffusion in solution, followed by adsorption at a different place) for the strongly bound or chemisorbed ligands. The probabilities for the different MC moves are adjusted during the equilibration period to achieve on average one accepted identity exchange move per cycle, and the remainder is divided in a 1:1:3 ratio between translational, rotational, and CBMC moves. The production periods consist of at least 500 000 MC cycles.

Most of the simulations are carried out for a single nanoparticle, but for the intermediate nanoparticle diameter, simulations are also performed for a system of two isolated nanoparticles (i.e., placed at a distance that ligands do not interact) that are in thermodynamic contact through identity exchange moves involving one ligand on each nanoparticle. In this case, initial configurations with either random spatial distributions of ligands but equimolar composition on both nanoparticles or with fully separated distributions (i.e., all long alkanethiolates placed on one nanoparticle and all short alkanethiolates on the other) are investigated.

All simulations are performed at room temperature (298 K) and without explicit solvent molecules, a setup that corresponds to a poor solvent (e.g., ethanol) as used in the experiments by Jackson et al. [238]. In all cases, four independent runs are carried out for each type of system (for the single nanoparticle systems, these are three different diameters, three different ligand mixtures, and two different initial arrangements of the ligands; for the two-nanoparticle systems, these are only the three different ligand mixtures and two different initial arrangements of the ligands).

The nonbonded interactions between different alkanethiolate molecules and beads

on the same ligand separated by more than three bonds are described by Lennard–Jones (LJ) potentials with the parameters for CH_x–CH_x and S–CH_x taken from the united-atom version of the TraPPE force field [252, 253]. The LJ σ parameter of 4.97 Å for S–S interactions [254] is specially adapted to yield the headgroup ordering observed for alkanethiolate SAMs, and hence, partial charges are not used for the alkanethiolates. The LJ parameters for Au are taken from previous work on gold nanoparticles [242]. The Lorentz–Berthelot combining rules [255] are used to determine LJ parameters for unlike nonbonded interactions. This does not include the strong S–Au interactions that are modeled using a Morse potential with parameters taken from Zhao et al. [256]. A cutoff distance at 14 Å is used for the nonbonded interactions and, because of the heterogeneous density of the isolated nanoparticles, tail corrections are not applied. The alkanethiolates are treated as semi-flexible with rigid bond length, but flexible bend and dihedral angles. The parameters for these bonded interactions are taken from the TraPPE–UA force field [252, 253].

7.3 Results and Discussion

7.3.1 Equilibration

One of the challenges in molecular simulation studies is determining whether the run length is sufficient so that the system under investigation has reached thermodynamic equilibrium, or whether the free energy landscape is too rugged and the system has found a kinetically stable, but thermodynamically metastable, basin in the free energy landscape. In order to assess the extent of the equilibration, the systems are initialized here with two very different configurations: an ordered Janus-like arrangement with ligands of different lengths occupying two different sides of the nanoparticle or a random distribution of the head group locations for ligands of different length. The approach to equilibrium is monitored by the average number of nearest neighbors of specific types (head groups belonging to either short or long ligands) because this property is found to change at the slowest rate, whereas the potential energy changes rapidly during the initial equilibration without much change in the spatial distribution of the ligands. Choosing a distance cut-off at 6.5 Å corresponding to the first minimum in the head–head radial distribution function (RDF), allows one to find the six nearest neighbor

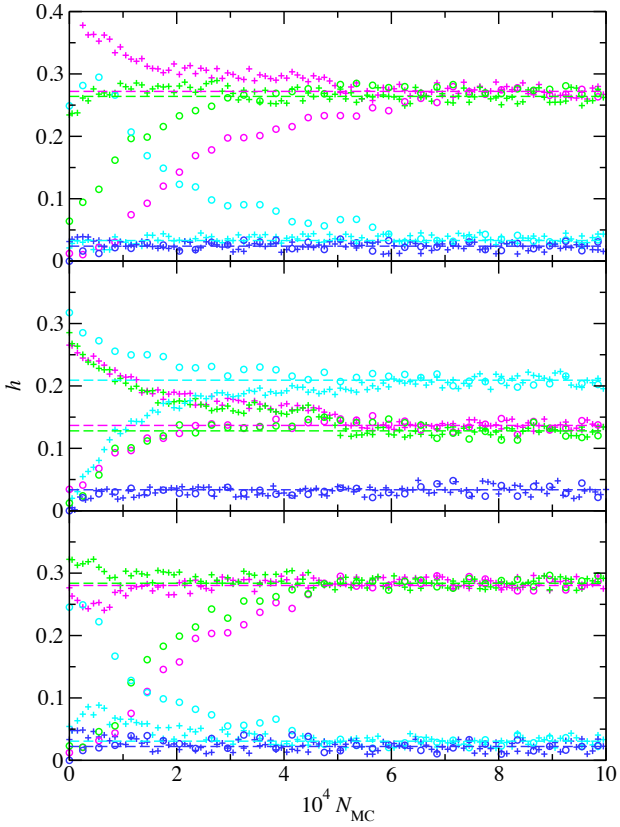

Figure 7.1: Evolution of the fraction of head groups being surrounded by a specific number of neighboring head groups belonging to short-chain ligands for the 70 Å nanoparticle covered with C6/C10 (top), C6/C14 (middle), and C10/C14 (bottom) mixtures. The trajectories start from the beginning of the equilibration period utilizing identity exchange moves; note that this equilibration period consists of 200 000 MC cycles, and only the first half is shown for clarity. The data show the average computed for blocks of 1000 MC cycles. The plusses and circles correspond to trajectories started from random and Janus-like initial configurations, respectively. The dashed lines show the average values obtained from the production runs. The magenta and green colors denote short and long ligands with three short-chain neighbors, respectively; the cyan and blue colors denote short and long ligands with six short-chain neighbors, respectively.

head groups surrounding a given head group.

The evolution of the block averages for head groups surrounded by either an equal number of short- and long-chain ligands or only short-chain ligands on the 70 Å nanoparticle is illustrated for the three mixtures in Figure 7.1. The trajectories shown start from

the beginning of the second equilibration period that utilizes identity exchange moves, and the very different initial values ($N_{\mathrm{MC}} = 0$) indicate that the pre-equilibration period using only translational, rotational, and conformational moves is not sufficient to yield converged results for the spatial distribution of ligands. Diffusion is very much hindered by the dense packing of the sulfur head groups. Once the identity exchange moves are introduced, however, the systems are found to equilibrate in fewer than 80 000 MC cycles to statistically indistinguishable spatial distributions irrespective of the initial arrangement. That is, the identity exchange moves allow for rapid equilibration of the difference in the excess chemical potentials between the short- and long-chain ligands. Similarly, short equilibration periods are also observed for the other two nanoparticle diameters.

The differences in the converged values (most pronounced for the fraction of short ligands surrounded by six short ligands; cyan symbols in Figure 7.1) for the C6/C14 mixture compared to the other two mixtures indicate that their structure and phase behavior differ markedly. For some of the simulations for the C6/C14 mixture started from a random initial arrangement of head groups, striped arrangements of the different ligands are observed during the early part of the equilibration but are quickly converted to a Janus-like arrangement (see below).

7.3.2 Single Nanoparticle Systems

Spatial Distribution

To analyze which factors play a role in the spatial distribution of mixed SAMs confined to a single gold nanoparticle, nine systems are investigated here that are formed by the combination of three different nanoparticle diameters and three different ligand mixtures. To this extent, the length scale is important and one needs to distinguish between (i) a random arrangement of ligands that shows a stochastic distribution of the nearest neighbors surrounding a given ligand, (ii) a local composition heterogeneity where the distribution of nearest neighbors deviates from a stochastic distribution and (usually) exhibits an enrichment of like ligands in the nearest neighbor shell that decays rapidly for farther outlying neighbor shells, (iii) the formation of more than two domains (e.g., stripes) that each extent over a couple of nearest-neighbor distances), and (iv)

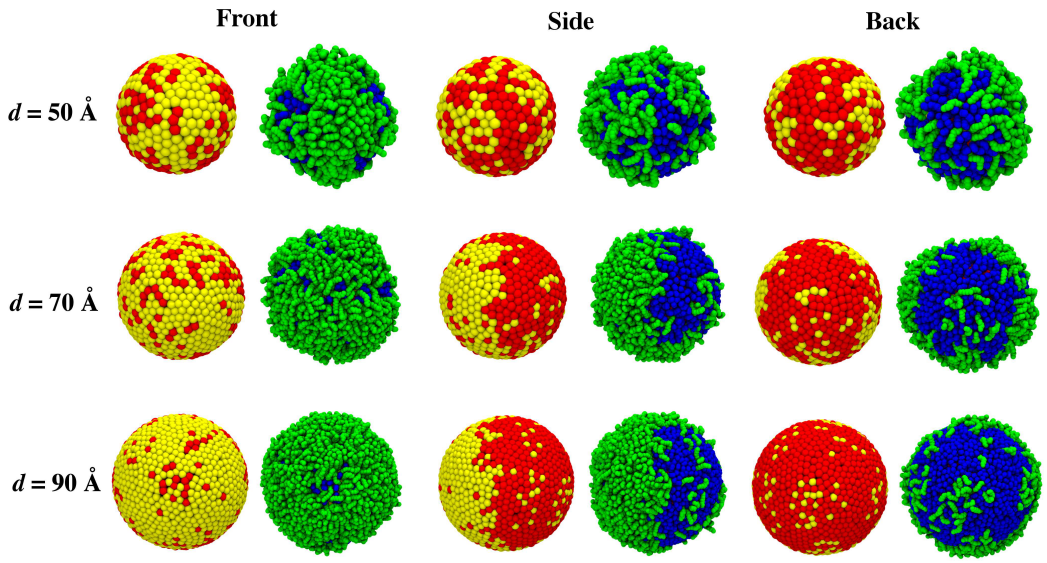

Figure 7.2: Snapshots of nanoparticles after equilibration for the C6/C14 mixture. Red and blue spheres represent the sulfur head group and CH_x units of the C6 ligand, respectively; yellow and green spheres represent sulfur head group and CH_x units of the C14 ligand, respectively. For each nanoparticle diameter, three orientations are shown that are selected according to the largest difference in local composition with the half enriched by C14 ligands denoted as the front.

phase separation that when confined to a single nanoparticle yields only two domains and, specifically, a Janus-like structure for equimolar mixtures.

Here, we start with a discussion of snapshots taken from equilibrated configurations for the nine systems (see Figures 7.2 and 7.3) because they provide visual guidance for subsequent quantitative analysis. For the C6/C14 mixture, we observe a Janus-like arrangement and phase separation for all three nanoparticle diameters. Moreover, the mutual miscibility appears to decrease with increasing nanoparticle diameter, but a tendency to aggregate persists for the minority species for all three nanoparticle diameters. These observations differ from previous MD studies [242], where the mutual miscibility was found to be extremely low irrespective whether Janus-like or striped arrangements are present. This difference can be explained by the treatment of CH_x–CH_x interactions that are treated here in the same manner irrespective whether the two units belong to ligands of the same type or of different types, whereas Gorai et al. [242] treat them as

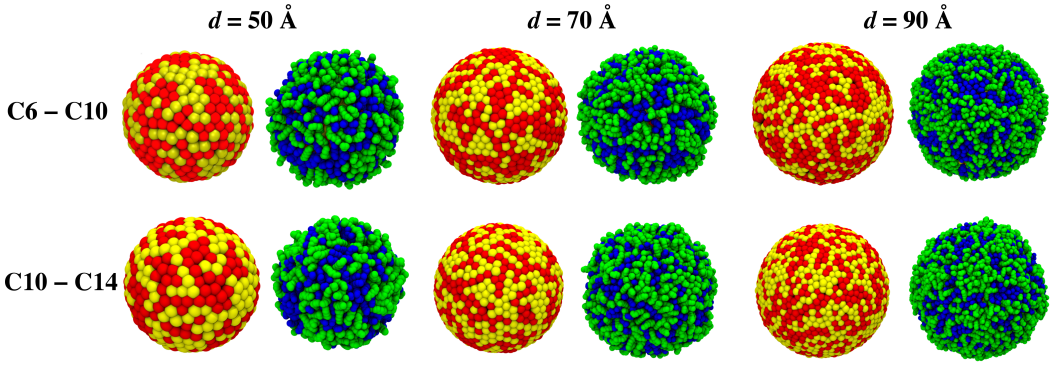

Figure 7.3: Snapshots of nanoparticles after equilibration for the C6/C10 and C10/C14 mixture. Red and blue spheres represent the sulfur head group and CH_x units of the shorter ligands, respectively; yellow and green spheres represent sulfur head group and CH_x units of the longer ligands, respectively.

purely repulsive for unlike pairs and with the LJ potential for like pairs. The latter leads to a much larger enthalpic driving force for demixing. Singh et al. [241] reported Janus-like arrangements for the smallest nanoparticle and an increasing number of stripes with increasing diameter for larger nanoparticles, whereas the current simulation yield an enhancement of the Janus-like arrangement (reduced mutual miscibility) with increasing diameter. The snapshots also provide initial evidence that the upper part of the longer ligands is significantly disordered with many conformational defects. Particularly, isolated longer ligands or those in small aggregates often fold over and adsorb on the surface provided by the shorter ligands.

Visual inspection of the nanoparticles coated with C6/C10 and C10/C14 ligand mixtures (see Figure 7.3) indicates strong local composition heterogeneities, but not the formation of ordered domains (e.g., stripes) or phase separation. Thus, phase separation into a Janus-like arrangement occurs only for ligand mixtures of sufficiently different lengths and, of course, the extent of the required length difference should depend on the temperature. There are no easily discernible differences caused by changes in the nanoparticle diameter. In the current model systems, it appears that the thermodynamic driving forces (i.e., the enthalpic gain from additional long–long chain interactions and the change in line tension from formation of larger domains) are not sufficient to favor ordered domains or Janus- like arrangements for ligands differing in length only

by four CH_2 units.

With regards to quantitative measures of the spatial distribution, we discuss first the distribution of ligands in the first solvation shell of the two types of ligands. Such data are represented here in compact form through the difference, $\Delta h(N_{\text{short}})$, in the average fraction of short ligands being surrounded by a specific number of short ligands and the average fraction of long ligands being surrounded by a specific number of short ligands (see Figure 7.4). In a random arrangement, this difference is zero because the average coordination would be the same for both ligands; of course there would be fluctuations around this average environment, and the instantaneous coordination would not be the same for all ligands. Local or global heterogeneities yield deviations from zero. In the present case, where like–like pairs are favored and an equimolar mixture is considered, the values of $\Delta h(N_{\text{short}})$ are negative for $N_{\text{short}} < 3$, positive for $N_{\text{short}} > 3$, and ≈ 0 for $N_{\text{short}} = 3$. Please note that the $\Delta h(N_{\text{short}}$ values do not need to be symmetric around $N_{\text{short}} = 3$ because a positive value for $N_{\text{short}} = 0$ reflects the arrangement in domains enriched by long ligands where long ligands are surrounded predominantly by long ligands, whereas $N_{\text{short}} = 0$ reflects the arrangement in domains enriched by the short chains.

The salient features of $\Delta h(N_{\text{short}})$ are the same for all three nanoparticle diameters, but the extent of the heterogeneities (the magnitude of the data points) increases with increasing nanoparticle diameter (see Figure 7.4). Decreased curvature leads to fewer disclinations [257], and in the limit of the infinite radius (i.e., a planar substrate), phase separation will occur [246, 247, 248].

Strong segregation is found for the C6/C14 mixture, and maxima in $|\Delta h(N_{\text{short}})|$ at $N_{\text{short}} = 0$ and 6 are observed for the 70 and 90 Å nanoparticles. For the other two ligand mixtures, the extrema in $\Delta h(N_{\text{short}})$ are found at intermediate N_{short} values. The local heterogeneity is somewhat more pronounced for the C6/C10 mixture than the C10/C14 mixture. This can likely be attributed to two factors: (i) the relative difference in chain length is about 20% larger for the C6/C10 mixture leading to a somewhat larger line tension, and (ii) the nanoparticle curvature results in more free volume for the upper part of the C10/C14 mixture and, hence, more conformational entropy.

To quantify the extent of the composition heterogeneity beyond the first neighbor

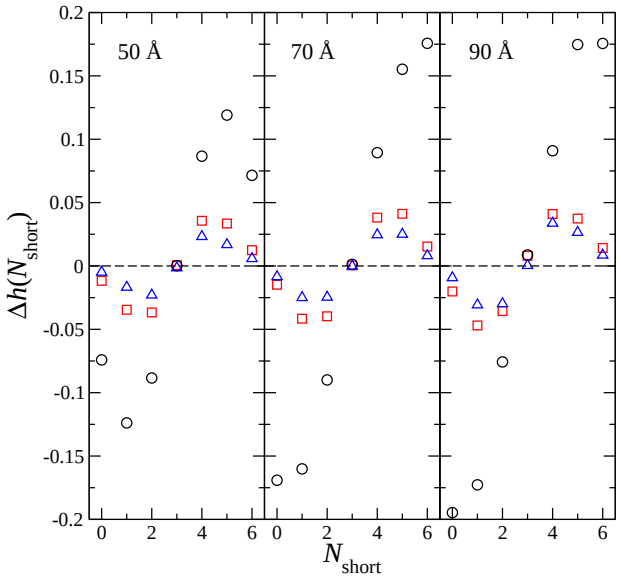

Figure 7.4: Difference in the average fractions of short and long ligands surrounded by a specific number of short ligands as nearest neighbor for the nanoparticles with diameters of 50 Å (left), 70 Å (middle), and 90 Å (right). The red squares, black circles, and blue triangles denote data for the C6/C10, C6/C14, and C10/C14 mixtures, respectively.

shell, the composition of the first 18 neighbors of each ligand is calculated (see Figure 7.5). For the C6/C10 and C10/C14 mixtures, these distributions are unimodal for all three nanoparticle diameters with the peak found at $S_{\rm short} = 9$. Since the nearest neighbor distribution discussed above yields local composition enrichments, this peak indicates the large fraction of molecules found at the interface of small domains that are enriched in the like component of the mixture. In a comparison between these two mixtures, one finds that the peak is slightly higher and narrower for the C10/C14 mixture, i.e., a larger fraction of ligands is present at interfaces or in disordered environments, in agreement with the smaller extent of the local demixing for this mixture compared to the C6/C10 mixture.

The distributions observed for the C6/C14 mixture exhibit qualitatively different features depending on the nanoparticle diameter. For the smallest nanoparticle, the distribution is unimodal but much broader than for the other two mixtures. The larger probabilities for the wings of the distribution indicate that more of the ligands are surrounded by like ligands as also observed in the snapshots. For the 50 Å nanoparticle,

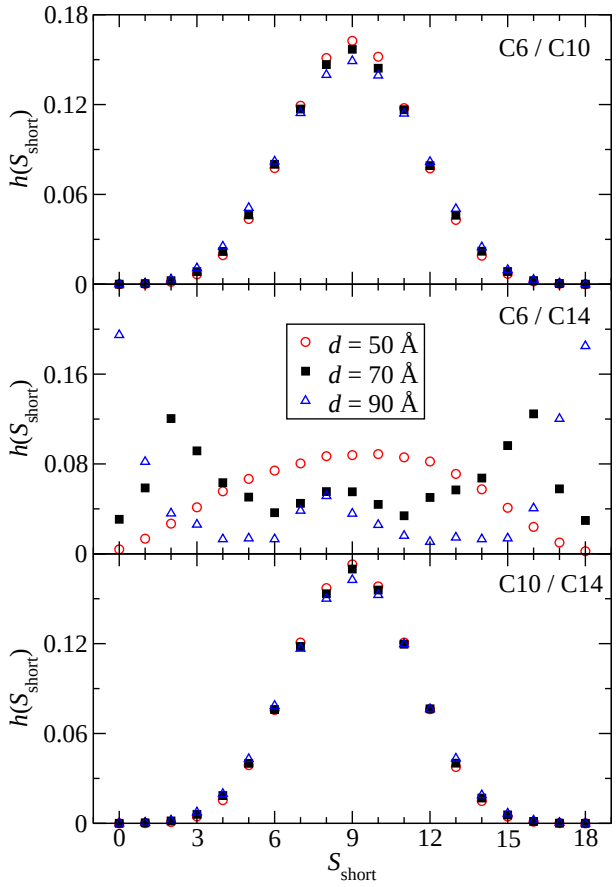

Figure 7.5: Distribution of ligands surrounded by a specific number of short ligands in the first two neighbor shells for the C6/C10 (top), C6/C14 (middle), and C10/C14 mixtures (bottom). The red circles, black filled squares, and blue triangles denote data for the nanoparticles with diameters of 50, 70, and 90 , respectively.

a short ligand surrounded by 18 short ligands would comprise about 9% of all short ligands. For the 70 Å nanoparticle, the distribution is trimodal with two larger peaks at $S_{\text{short}} = 3$ and 16 and a minor peak centered at $S_{\text{short}} = 9$ and 10. The latter reflects ligands at the interface of larger domains, whereas the former peaks are for ligands within a domain of like chains. The trimodal distribution is a good signature for the Janus-like arrangement that is so clearly visible in the snapshots (see Figure 7.2). For the 90 Å nanoparticle, the trimodal character becomes more pronounced with the two main peaks found at $S_{\text{short}} = 0$ and 18, i.e., indicating that the mutual miscibility

is relatively low. For an ordered arrangement of stripes, one would find a trimodal distribution only when the stripe width is larger than 5 ligands, and the distribution should not change significantly with the nanoparticle diameter.

Structural Analysis

One of the most widely used structural properties for alkyl thiol chains adsorbed onto metal surfaces is their tilt angle. The tilt angle of a ligand is computed here in analogy with those for SAMs on a flat substrate [254] as the angle between the surface normal vector passing through the sulfur head group and the vector connecting the corresponding sulfur head group and the terminal methyl group of the ligand. A more local property is the local bond order parameter that depends on the orientation of the surface normal with the vector formed by two CH_x groups separated by two bonds [254]. Velachi et al. [244] found no significant differences of the tilt angles for SAMs on nanoparticles regardless of wether the ligands are in a random, striped, or Janus-like arrangement, but they used a different definition of the tilt angle [258]. Tilt angles calculated for the systems studied in this work are summarized in Table 7.1. In general, the tilt angle increases with an increasing ligand length and decreases with an increasing nanoparticle diameter. Both of these trends agree with previous studies [258, 244], and are due to curvature effects that provide more free volume for the outer part of the SAMs. In addition, the curvature allows for a relatively broad distribution of tilt angles (as indicated here by the standard deviation; note that the standard error of the mean computed from the four independent trajectories for each system is smaller than 2° in all cases) and average tilt angles show little effect of whether a ligand of a given length is the short or long component or whether larger domains of like ligands are present.

Examination of radial bead density profiles provides further insight into the structure of the chemisorbed alkyl thiol chains (see Figure 7.6). For all systems, the lower part of the alkyl chain region is relatively well-ordered with about five clearly discernible peaks. For the C6/C10 and C10/C14 SAMs that are not found in a Janus-like arrangement, the radial density profile for the 50 Å nanoparticle exhibits considerably less order as indicated by the reduced magnitude of the density oscillations. In all cases, a gradual decay is observed for the outer part of the SAMs that is populated only by the longer ligands, i.e., these regions contain alkyl tails with significant conformational disorder.

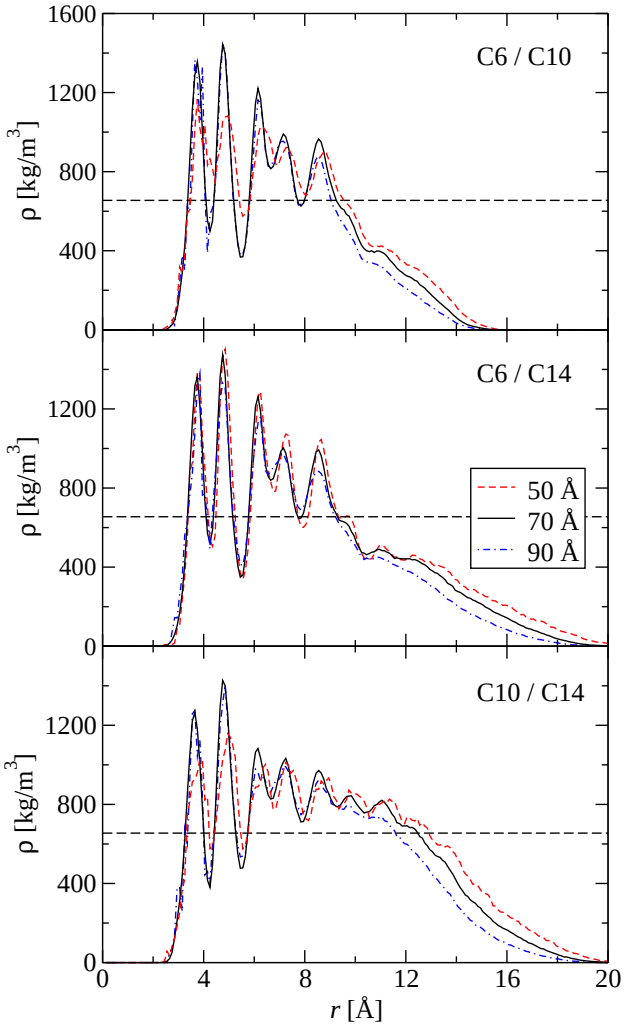

Figure 7.6: Radial density profiles of CH_x groups from the surface of the nanoparticle for the C6/C10 (top); C6/C14 (middle), and C10/C14 mixtures (bottom). The red dashed, black solid, and blue dash-dotted lines show data for the 50, 70, and 90 Å nanoparticles, respectively. The horizontal dashed line represents the density of bulk liquid *n*-hexane.

The Janus-like arrangement with less backfolding of chain backbones for the C6/C14 SAMs leads to an extension of the radial bead density profiles; i.e., these ligands are conformationally more ordered than would be expected for a spatial distribution with only local heterogeneities.

Table 7.1: Average Tilt Angles (in Degrees) for the Different Systems. [a]

	C6/C10		C6/C14		C10/C14	
d [Å]	C6	C10	C6	C14	C10	C14
50	23_{10}	32_{13}	24_{11}	40_{15}	29_{13}	39_{15}
70	21_{10}	27_{12}	21_{10}	35_{14}	26_{12}	35_{14}
90	19_{10}	24_{13}	20_{10}	30_{14}	22_{12}	30_{14}

[a] Numbers in Subscripts Represent the Standard Deviation.

7.3.3 Two Nanoparticle Systems

In experiments on SAM-coated nanoparticles, the solution contains numerous nanoparticles and the ligands in a mixture could, in principle, not only distribute spatially on a given nanoparticle but could also distribute between multiple nanoparticles. To mimic such an exchange of ligands between nanoparticles, we investigate a system of two nanoparticles where ligands can exchange. That is, the number of ligands per nanoparticle and the overall (equimolar) composition are fixed, but the ligands can exchange according to thermodynamic driving forces to yield two nanoparticles with different composition. In this case, only nanoparticles with a 70 Å diameter with a total of 812 ligands per nanoparticle are investigated. For the C6/C10 and C10/C14 mixtures, we find that this additional flexibility does not yield different compositions for the two nanoparticles, and the differences in the number of short ligands between the two nanoparticles are 3 ± 6 and 4 ± 6, respectively. Correspondingly, the spatial distributions and structural properties are indistinguishable for the one- and two-nanoparticle systems for the C6/C10 and C10/C14 mixtures. For the C6/C14 mixture, however, the composition heterogeneity is found to extend beyond the Janus-like arrangement to a composition difference of 124 ± 16 ligands between the two nanoparticles, i.e., a "phase" separation with one C6-rich and one C14-rich nanoparticle and a mole fraction of 0.576 ± 0.010.

7.4 Conclusions

In the current simulations for nanoparticles coated with SAMs consisting of an equimolar mixture of alkanethiolates differing only in their chain length (C6, C10, and C14), we observe, at room temperature, local composition heterogeneities for SAMs formed by ligand mixtures with a length difference of four methylene groups. In contrast, the C6/C14 mixture yields a Janus-like arrangement when confined to a single nanoparticle. For the latter mixture, a striped arrangement is only transient before the Janus-like arrangement is achieved. Increasing the nanoparticle diameter leads to increased conformational ordering and composition heterogeneity.

References

[1] McQuarie, D. A. *Statistical Mechanics;* University Science Books: Mill Valley, CA, 2000.

[2] Tuckerman, M. E. *Statistical Mechanics: Theory and Molecular Simulation;* Oxford University Press: New York, 2010.

[3] Frenkel, D.; Smit, B. *Understanding Molecular Simulation, Second Edition: From Algorithms to Applications (Computational Science);* Academic Press: Cambrdige, 2nd ed.; 2001.

[4] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids;* Clarendon Press: Oxford, U.K., 1987.

[5] Shaw, M. S. *J. Chem. Phys.* **1991,** *94,* 7550–7553.

[6] Smith, W. R.; Triska, B. *J. Chem. Phys.* **1994,** *100,* 3019–3027.

[7] Johnson, J. K.; Panagiotopoulos, A. Z.; Gubbins, K. E. *Mol. Phys.* **1994,** *81,* 717–733.

[8] Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953,** *21,* 1087–1092.

[9] Busch, N. A.; Wertheim, M. S.; Yarmush, M. L. *J. Chem. Phys.* **1996,** *104,* 3962–3975.

[10] Tsangaris, D. M.; de Pablo, J. J. *J. Chem. Phys.* **1994,** *101,* 1477–1489.

[11] Visco, D. P. J.; Kofke, D. A. *J. Chem. Phys.* **1999,** *110,* 5493–5502.

[12] Chen, B.; Siepmann, J. I. *J. Phys. Chem. B* **2000,** *104,* 8725–8734.

[13] Chen, B.; Siepmann, J. I. *J. Phys. Chem. B* **2001,** *105,* 11275–11282.

[14] Busch, N. A.; Wertheim, M. S.; Yarmush, M. L. *J. Chem. Phys.* **1996,** *104,* 3962–3975.

[15] Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953,** *21,* 1087–1092.

[16] Wierzchowski, S.; Kofke, D. A. *J. Chem. Phys.* **2001,** *114,* 8752–8762.

[17] Panagiotopoulos, A. Z. *Mol. Phys.* **1987,** *61,* 37–41.

[18] Panagiotopoulos, A. Z.; Quirke, N.; Stapleton, M.; Tildesley, D. J. *Mol. Phys.* **1988,** *63,* 527–545.

[19] Chen, B.; Siepmann, J. I.; Oh, K. J.; Klein, M. L. *J. Chem. Phys.* **2001,** *115,* 10903–10913.

[20] Chen, B.; Siepmann, J. I.; Klein, M. L. *J. Phys. Chem. A* **2005,** *109,* 1137–1145.

[21] Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. USA* **2002,** *99,* 12562–12566.

[22] Zwanzig, R. W. *J. Chem. Phys.* **1954,** *22,* 1420–1426.

[23] Calvo, F. *Mol. Phys.* **2002,** *100,* 3421–3427.

[24] Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989,** *18,* 431–492.

[25] Torrie, G.; Valleau, J. *J. Comput. Phys.* **1977,** *23,* 187–199.

[26] Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992,** *13,* 1011–1021.

[27] Siepmann, J. I.; McDonald, I. R. *Mol. Phys.* **1992,** *75,* 255–259.

[28] Siepmann, J. I. *Mol. Phys.* **1990,** *70,* 1145-1158.

[29] Siepmann, J. I.; Frenkel, D. *Mol. Phys.* **1992,** *75,* 59-70.

[30] Rafferty, J. L.; Siepmann, J. I.; Schure, M. R. *J. Chromatogr. A* **2011,** *1218,* 9183–9193.

[31] P. Chen, Q.; D. Chu, J.; F. DeJaco, R.; P. Lodge, T.; Ilja Siepmann, J. **2016,** *49,* 3975–3985.

[32] Fetisov, E. O.; Kuo, I.-F. W.; Knight, C.; VandeVondele, J.; Van Voorhis, T.; Siepmann, J. I. *ACS Cent. Sci.* **2016,** *2,* 409–415.

[33] Santiso, E. E.; Gubbins, K. E. *Mol. Simul.* **2004,** *30,* 699–748.

[34] van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. *J. Phys. Chem. A* **2001,** *105,* 9396–9409.

[35] Brenner, D. W.; A., S. O.; Harrison, J. A.; Stuart, S. J.; Ni, B.; Sinnott, S. B. *J. Phys.: Condens. Matter* **2002,** *14,* 783–802.

[36] Farah, K.; Müller-Plathe, F.; Böhm, M. C. *ChemPhysChem* **2012,** *13,* 1127–1151.

[37] Turner, H. C.; Brennan, J. K.; Lísal, M.; Smith, W. R.; Johnson, K. J.; Gubbins, K. E. *Mol. Simul.* **2008,** *34,* 119–146.

[38] Leigh, G. J. *The World's Greatest Fix: A History of Nitrogen and Agriculture;* Oxford University Press: New York, 2004.

[39] Fisher, K.; Newton, W. E. *Nitrogen Fixation at the Millennium;* Elsevier: Amsterdam, The Netherlands, 2002.

[40] Tomeczek, J.; Gradoń, B. *Combust. Sci. Tech.* **1997,** *125,* 159–180.

[41] Goldman, N.; Bastea, S. *J. Phys. Chem. A* **2014,** *118,* 2897–2903.

[42] Seshadri, D. N.; Viswanath, D. S.; Kuloor, N. R. *AIChE J.* **1970,** *16,* 420–425.

[43] Quintana-Lacaci, G.; Agúndez, M.; Cernicharo, J.; Bujarrabal, V.; Sanchez Contreras, C.; Castro-Carrizo, A.; Alcolea, J. *Astron. Astrophys.* **2013,** *560,* L2.

[44] Chase, M. W. J. *NIST-JANAF Thermochemical Tables;* Journal of Physical and Chemical Reference Data Monographs American Institute of Physics: New York, 1998.

[45] Jones, R. O. *Rev. Mod. Phys.* **2015,** *87,* 897–923.

[46] Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1999,** *103,* 4508–4517.

[47] Hutter, J.; Ianuzzi, M.; Schiffmann, F.; VandeVondele, J. *WIREs: Comput. Mol. Sci.* **2014,** *4,* 15–25.

[48] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Comm.* **2005,** *167,* 103–128.

[49] Becke, A. D. *Phys. Rev. A* **1988,** *38,* 3098–3100.

[50] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988,** *37,* 785–789.

[51] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010,** *132,* 154104.

[52] VandeVondele, J.; Hutter, J. *J. Chem. Phys.* **2007,** *127,* 114105.

[53] Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996,** *54,* 1703–1710.

[54] Hartwingsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998,** *58,* 3641–3662.

[55] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008,** *120,* 215–241.

[56] Vydrov, O. A.; Van Voorhis, T. *J. Chem. Phys.* **2010,** *132,* 164113.

[57] Sabatini, R.; Gorni, T.; de Gironcoli, S. *Phys. Rev. B* **2013,** *87,* 041108.

[58] Bouanich, J.-P. *J. Quant. Spectrosc. Radiat. Transfer* **1992,** *47,* 243-250.

[59] Stillinger, F. H. *J. Chem. Phys.* **1963,** *38,* 1486–1494.

[60] Kuo, I. F. W.; Mundy, C. J.; J., M. M.; Siepmann, J. I.; VandeVondele, J.; Sprik, M.; Hutter, J.; Chen, B.; Klein, M. L.; Mohamed, F.; Krack, M.; Parrinello, M. *J. Phys. Chem. B* **2004,** *108,* 12990–12998.

[61] McGrath, M. J.; Siepmann, J. I.; Kuo, I.-F. W.; Mundy, C. J.; VandeVondele, J.; Hutter, J.; Mohamed, F.; Krack, M. *J. Phys. Chem. A.* **2006,** *110,* 640–646.

[62] Saitta, A. M.; Saija, F. *Proc. Natl. Acad. Sci. USA* **2014,** *111,* 13768–13773.

[63] Wang, L.-P.; Titov, A.; McGibbon, R.; Liu, F.; Pande, V. S.; Martínez, T. J. *Nat. Chem.* **2014,** *6,* 1044–1048.

[64] Bastea, S.; Fried, L. E. *J. Chem. Phys.* **2008,** *128,* 174502.

[65] Bastea, S.; Fried, L. E. Chemical Equilibrium Detonation. In *Shock Waves Science and Technology Library, Vol. 6: Detonation Dynamics*; Zhang, F., Ed.; Springer: Berlin, Heidelberg, 2012.

[66] Schott, G. L.; Shaw, M. S.; Johnson, J. D. *J. Chem. Phys.* **1985,** *82,* 4264–4275.

[67] Gijsbertsen, A.; Siu, W.; Kling, M. F.; Johnsson, P.; Jansen, P.; Stolte, S.; Vrakking, M. J. J. *Phys. Rev. Lett.* **2007,** *99,* 213003.

[68] Zen, A.; Trout, B. L.; Guidoni, L. *J. Chem. Phys.* **2014,** *141,* 014305.

[69] McGrath, M. J.; Siepmann, J. I.; Kuo, I.-F. W.; Mundy, C. J. *Mol. Phys.* **2006,** *104,* 3619–3626.

[70] Chen, B.; Potoff, J. J.; Siepmann, J. I. *J. Phys. Chem. B* **2001,** *105,* 3093–3104.

[71] Fetisov, E. O.; Shah, M. S.; Knight, C.; Tsapatsis, M.; Siepmann, J. I. *ChemPhysChem* **2018,** *19,* 512–518.

[72] Sada, E.; Kumazawa, H.; Butt, M.; Hayashi, D. *Chem. Eng. Sci.* **1976,** *31,* 839–841.

[73] Rivas, O. R.; Prausnitz, J. M. *AIChE J.* **1979,** *25,* 975–984.

[74] Heintz, Y. J.; Sehabiague, L.; Morsi, B. I.; Jones, K. L.; Luebke, D. R.; Pennline, H. W. *Energy Fuel.* **2009,** *23,* 4822–4830.

[75] Lei, Z.; Dai, C.; Chen, B. *Chem. Rev.* **2014,** *114,* 1289–1326.

[76] Warrag, S. E.; Peters, C. J.; Kroon, M. C. *Curr. Opin. Green Sustain. Chem.* **2017,** *5,* 55–60.

[77] Crespo, D.; Qi, G.; Wang, Y.; Yang, F. H.; Yang, R. T. *Ind. Eng. Chem. Res.* **2008,** *47,* 1238–1244.

[78] Kumar, P.; Sung, C.-Y.; Muraza, O.; Cococcioni, M.; Hashimi, S. A.; McCormick, A.; Tsapatsis, M. *Microporous and Mesoporous Mater.* **2011,** *146,* 127–133.

[79] Shah, M. S.; Tsapatsis, M.; Siepmann, J. I. *Angew. Chem. Int. Ed.* **2016,** *55,* 5938–5942.

[80] Shah, M. S.; Tsapatsis, M.; Siepmann, J. I. *Chem. Rev.* **2017,** *117,* 9755–9803.

[81] Bülow, M.; Lutz, W.; Suckow, M. *Stud. Surf. Sci. Catal.* **1998,** *120,* 301–345.

[82] Fellmuth, P.; Lutz, W.; Bülow, M. *Zeolites* **1987,** *7,* 367–371.

[83] Bülow, M.; Micke, A. Utilization of Zeolites for Processes of Gas Desulphurization. In *Fundamentals of Adsorption: Proceedings of the Fifth International Conference on Fundamentals of Adsorption*; LeVan, M. D., Ed.; Springer US: Boston, MA, 1996.

[84] Collins, J. J. *Oil Gas J.* **1963,** *270,* 97–99.

[85] Lutz, W.; Suckow, M.; Bülow, M. *Gas Sep. Purif.* **1990,** *4,* 190–196.

[86] Higgins, J.; LaPierre, R.; Schlenker, J.; Rohrman, A.; Wood, J.; Kerr, G.; Rohrbaugh, W. *Zeolites* **1988,** *8,* 446–452.

[87] Lísal, M.; Nezbeda, I.; Smith, W. R. *J. Chem. Phys.* **1999,** *110,* 8597–8604.

[88] Lísal, M.; Smith, W. R.; Nezbeda, I. *AIChE J.* **2000,** *46,* 866–875.

[89] Perea, D. E.; Arslan, I.; Liu, J.; Ristanović, Z.; Kovarik, L.; Arey, B. W.; Lercher, J. A.; Bare, S. R.; Weckhuysen, B. M. *Nat. Commun.* **2015,** *6,* 7589.

[90] W. Löwenstein, *Am. Mineral.* **1954,** *39,* 92–96.

[91] Baerlocher, C.; McCusker, L. Database of Zeolite Structures: http://www.iza-structure.org/databases.

[92] McGrath, M. J.; Siepmann, J. I.; Kuo, I.-F. W.; Mundy, C. J.; VandeVondele, J.; Hutter, J.; Mohamed, F.; Krack, M. *ChemPhysChem* **2005,** *6,* 18941901.

[93] Bai, P.; Tsapatsis, M.; Siepmann, J. I. *J. Phys. Chem. C* **2013,** *117,* 24375–24387.

[94] Shah, M. S.; Tsapatsis, M.; Siepmann, J. I. *J. Phys. Chem. B* **2015,** *119,* 7041–7052.

[95] Potoff, J. J.; Siepmann, J. I. *AIChE J.* **2001,** *47,* 1676–1682.

[96] Hutter, J.; Ianuzzi, M.; Schiffmann, F.; VandeVondele, J. *WIREs: Comput. Mol. Sci.* **2014,** *4,* 15–25.

[97] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996,** *77,* 3865–3868.

[98] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010,** *132,* 154104.

[99] Fischer, M.; Evers, F. O.; Formalik, F.; Olejniczak, A. *Theor. Chem. Acc.* **2016,** *135,* 257.

[100] Frisch, M. J. *et al.* "Gaussian 09 Revision D.01", 2009 Gaussian Inc. Wallingford CT.

[101] Orkoulas, G.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **1994,** *101,* 1452–1459.

[102] Duane, S.; Kennedy, A.; Pendleton, B. J.; Roweth, D. *Phys. Lett. B* **1987,** *195,* 216–222.

[103] Turner, C. H.; Johnson, J. K.; Gubbins, K. E. *J. Chem. Phys.* **2001,** *114,* 1851–1859.

[104] Lísal, M.; Cosoli, P.; Smith, W. R.; Jain, S. K.; Gubbins, K. E. *Fluid Phase Equilib.* **2008,** *272,* 18–31.

[105] Lísal, M.; Brennan, J. K.; Smith, W. R. *J. Chem. Phys.* **2006,** *110,* 064712.

[106] Turner, C. H.; Gubbins, K. E. *J. Chem. Phys.* **2003,** *119,* 6057–6067.

[107] Hansen, N.; Jakobtorweihen, S.; Keil, F. J. *J. Chem. Phys.* **2005,** *122,* 164705.

[108] Mullen, R. G.; Maginn, E. J. *J. Chem. Theory Comput.* **2017,** *13,* 4054–4062.

[109] Henzler, K.; Fetisov, E. O.; Galib, M.; Baer, M. D.; Legg, B. A.; Borca, C.; Xto, J. M.; Pin, S.; Fulton, J. L.; Schenter, G. K.; Govind, N.; Siepmann, J. I.; Mundy, C. J.; Huthwelker, T.; De Yoreo, J. J. *Sci. Adv.* **2018,** *4,* eaao6283.

[110] Kashchiev, D. *J. Chem. Phys.* **2006,** *125,* 014502.

[111] Gebauer, D.; Völkel, A.; Cölfen, H. *Science* **2008,** *322,* 1819–1822.

[112] Gebauer, D.; Cölfen, H. *Nano Today* **2011,** *6,* 564–584.

[113] Demichelis, R.; Raiteri, P.; Gale, J. D.; Quigley, D.; Gebauer, D. *Nat. Commun.* **2011,** *2,* 590.

[114] Wallace, A. F.; Hedges, L. O.; Fernandez-Martinez, A.; Raiteri, P.; Gale, J. D.; Waychunas, G. A.; Whitelam, S.; Banfield, J. F.; De Yoreo, J. J. *Science* **2013,** *341,* 885–889.

[115] Kellermeier, M.; Raiteri, P.; Berg, J. K.; Kempter, A.; Gale, J. D.; Gebauer, D. *ChemPhysChem* **2016,** *17,* 3535–3541.

[116] Pouget, E. M.; Bomans, P. H. H.; Goos, J. A. C. M.; Frederik, P. M.; de With, G.; Sommerdijk, N. A. J. M. *Science* **2009,** *323,* 1455–1458.

[117] Plummer, L. N.; Busenberg, E. *Geochim. Cosmochim. Acta* **1982,** *46,* 1011–1040.

[118] Smeets, P. J. M.; Finney, A. R.; Habraken, W. J. E. M.; Nudelman, F.; Friedrich, H.; Laven, J.; De Yoreo, J. J.; Rodger, P. M.; Sommerdijk, N. A. J. M. *Proc. Natl. Acad. Sci. USA* **2017,** *114,* E7882–E7890.

[119] Tribello, G. A.; Bruneval, F.; Liew, C.; Parrinello, M. *J. Phys. Chem. B* **2009,** *113,* 11680–11687.

[120] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965,** *140,* A1133–A1138.

[121] Lippert, G.; Hutter, J.; Parinello, M. *Mol. Phys.* **1997,** *92,* 477–488.

[122] Galib, M.; Baer, M. D.; Skinner, L. B.; Mundy, C. J.; Huthwelker, T.; Schenter, G. K.; Benmore, C. J.; Govind, N.; Fulton, J. L. *J. Chem. Phys.* **2017,** *146,* 084504.

[123] Bogatko, S.; Caut, E.; Bylaska, E.; Schenter, G.; Fulton, J.; Weare, J. *Chem. Eur. J.* **2013,** *19,* 3047–3060.

[124] Raiteri, P.; Gale, J. D.; Quigley, D.; Rodger, P. M. *J. Phys. Chem. C* **2010,** *114,* 5997–6010.

[125] Raiteri, P.; Gale, J. D. *J. Am. Chem. Soc.* **2010,** *132,* 17623–17634.

[126] Baer, M. D.; Mundy, C. J. *J. Phys. Chem. B* **2016,** *120,* 1885–1893.

[127] Raiteri, P.; Demichelis, R.; Gale, J. D. *J. Phys. Chem. C* **2015,** *119,* 24447–24458.

[128] Legg, B. A.; Yoreo, J. J. D. *J. Chem. Phys.* **2016,** *145,* 211921.

[129] Duignan, T. T.; Baer, M. D.; Mundy, C. J. *Curr. Opin. Colloid Interface Sci.* **2016,** *23,* 58–65.

[130] Hale, B. N. *Aust. J. Phys.* **1996,** *49,* 425–434.

[131] Zou, Z.; Habraken, W. J. E. M.; Bertinetti, L.; Politi, Y.; Gal, A.; Weiner, S.; Addadi, L.; Fratzl, P. *Adv. Mater. Interfaces* **2017,** *4,* 1600076.

[132] Stöhr, J. *NEXAFS Spectroscopy;* Springer-Verlag: Heidelberg, Germany, 1992.

[133] Rehr, J.; Ankudinov, A. *Coord. Chem. Rev.* **2005,** *249,* 131–140.

[134] Brown, M. A.; Huthwelker, T.; Beloqui Redondo, A.; Janousch, M.; Faubel, M.; Arrell, C. A.; Scarongella, M.; Chergui, M.; van Bokhoven, J. A. *J. Phys. Chem. Lett.* **2012,** *3,* 231–235.

[135] Pin, S.; Huthwelker, T.; Brown, M. A.; Vogel, F. *J. Phys. Chem. A* **2013,** *117,* 8368–8376.

[136] Flank, A.-M.; Cauchon, G.; Lagarde, P.; Bac, S.; Janousch, M.; Wetter, R.; Dubuisson, J.-M.; Idir, M.; Langlois, F.; Moreno, T.; Vantelon, D. *Nucl. Instrum. Methods Phys. Res., Sect. B* **2006,** *246,* 269–274.

[137] Fulton, J. L.; Heald, S. M.; Badyal, Y. S.; Simonson, J. M. *J. Phys. Chem. A* **2003,** *107,* 4688–4696.

[138] Michel, F. M.; MacDonald, J.; Feng, J.; Phillips, B. L.; Ehm, L.; Tarabrella, C.; Parise, J. B.; Reeder, R. J. *Chem. Mater.* **2008,** *20,* 4720–4728.

[139] Lopata, K.; Van Kuiken, B. E.; Khalil, M.; Govind, N. *J. Chem. Theory Comput.* **2012,** *8,* 3284–3292.

[140] Fulton, J. L.; Govind, N.; Huthwelker, T.; Bylaska, E. J.; Vjunov, A.; Pin, S.; Smurthwaite, T. D. *J. Phys. Chem. B* **2015,** *119,* 8380–8388.

[141] Martin-Diaconescu, V.; Gennari, M.; Gerey, B.; Tsui, E.; Kanady, J.; Tran, R.; Pécaut, J.; Maganas, D.; Krewald, V.; Gouré, E.; Duboc, C.; Yano, J.; Agapie, T.; Collomb, M.-N.; DeBeer, S. *Inorg. Chem.* **2015,** *54,* 1283–1292.

[142] Kellermeier, M.; Picker, A.; Kempter, A.; Cölfen, H.; Gebauer, D. *Adv. Mater.* **2014,** *26,* 752–757.

[143] Genovese, D.; Montalti, M.; Otálora, F.; Gómez-Morales, J.; Sancho-Tomás, M.; Falini, G.; García-Ruiz, J. M. *Crys. Growth Des.* **2016,** *16,* 4173–4177.

[144] Andersson, M. P.; Dobberschütz, S.; Sand, K. K.; Tobler, D. J.; DeYoreo, J. J.; Stipp, S. L. S. *Angew. Chem. Int. Ed.* **2016,** *55,* 11086–11090.

[145] Carino, A.; Testino, A.; Andalibi, M. R.; Pilger, F.; Bowen, P.; Ludwig, C. *Crys. Growth Des.* **2017,** *17,* 2006–2015.

[146] Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys.* **1992,** *97,* 2635–2643.

[147] Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996,** *54,* 1703–1710.

[148] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988,** *37,* 785–789.

[149] Grimme, S. *J. Comput. Chem.* **2006,** *27,* 1787–1799.

[150] Wood, W. W.; Parker, F. R. *J. Chem. Phys.* **1957,** *27,* 720–733.

[151] Fetisov, E. O.; Harwood, D. B.; Kuo, I.-F. W.; Warrag, S. E. E.; Kroon, M. C.; Peters, C. J.; Siepmann, J. I. *J. Phys. Chem. B* **2018,** *122,* 1245–1254.

[152] Welton, T. *Chem. Rev.* **1999,** *99,* 2071–2084.

[153] Brennecke, J. F.; Maginn, E. J. *AIChE J.* **2001,** *47,* 2384–2389.

[154] Plechkova, N. V.; Seddon, K. R. *Chem. Soc. Rev.* **2008,** *37,* 123–150.

[155] MacFarlane, D. R.; Tachikawa, N.; Forsyth, M.; Pringle, J. M.; Howlett, P. C.; Elliott, G. D.; Davis, J. H.; Watanabe, M.; Simon, P.; Angell, C. A. *Energy Environ. Sci.* **2014,** *7,* 232–250.

[156] Fredlake, C. P.; Crosthwaite, J. M.; Hert, D. G.; Aki, S. N. V. K.; Brennecke, J. F. *J. Chem. Eng. Data* **2004,** *49,* 954–964.

[157] Smith, E. L.; Abbott, A. P.; Ryder, K. S. *Chem. Rev.* **2014,** *114,* 11060–11082.

[158] Abbott, A. P.; Capper, G.; Davies, D. L.; Rasheed, R. K.; Tambyrajah, V. *Chem. Commun.* **2003,** 70–71.

[159] Zhang, Q.; Vigier, K. D. O.; Royer, S.; Jérôme, F. *Chem. Soc. Rev.* **2012,** *41,* 7108–7146.

[160] Francisco, M.; van den Bruinhorst, A.; Kroon, M. C. *Angew. Chem. Int. Ed.* **2013,** *52,* 3074–3085.

[161] Parnham, E. R.; Morris, R. E. *Acc. Chem. Res.* **2007,** *40,* 1005–1013.

[162] Abbott, A. P.; Capper, G.; McKenzie, K. J.; Ryder, K. S. *J. Electroanal. Chem.* **2007,** *599,* 288–294.

[163] Malaquias, J. C.; Steichen, M.; Thomassey, M.; Dale, P. J. *Electrochim. Acta* **2013,** *103,* 15–22.

[164] Zhao, H.; Baker, G. A. *J. Chem. Technol. Biotechnol.* **2013,** *88,* 3–12.

[165] Osch, D. J. G. P. v.; Zubeir, L. F.; Bruinhorst, A. v. d.; Rocha, M. A. A.; Kroon, M. C. *Green Chem.* **2015,** *17,* 4518–4521.

[166] Dai, Y.; Witkamp, G.-J.; Verpoorte, R.; Choi, Y. H. *Food Chem.* **2015,** *187,* 14–19.

[167] Francisco, M.; Bruinhorst, A. v. d.; Zubeir, L. F.; Peters, C. J.; Kroon, M. C. *Fluid Phase Equilib.* **2013,** *340,* 77–84.

[168] García, G.; Atilhan, M.; Aparicio, S. *Chem. Phys. Lett.* **2015,** *634,* 151–155.

[169] Hammond, O. S.; Bowron, D. T.; Edler, K. J. *Green Chem.* **2016,** *18,* 2736–2744.

[170] Zahn, S.; Kirchner, B.; Mollenhauer, D. *ChemPhysChem* **2016,** *17,* 3354–3358.

[171] Perkins, S. L.; Painter, P.; Colina, C. M. *J. Phys. Chem. B* **2013,** *117,* 10250–10260.

[172] Perkins, S. L.; Painter, P.; Colina, C. M. *J. Chem. Eng. Data* **2014,** *59,* 10250–10260.

[173] Sun, H.; Li, Y.; Wu, X.; Li, G. *J. Mol. Model* **2013,** *19,* 2433–2441.

[174] Shah, D.; Mjalli, F. S. *Phys. Chem. Chem. Phys.* **2014,** *16,* 23900–23907.

[175] García, G.; Atilhan, M.; Aparicio, S. *J. Molec. Liq.* **2015,** *211,* 506–514.

[176] Mainberger, S.; Kindlein, M.; Bezold, F.; Elts, E.; Minceva, M.; Briesen, H. *Mol. Phys.* **2017,** *115,* 1309–1321.

[177] Del Popolo, M. G.; Lynden-Bell, R. M.; Kohanoff, J. *J. Phys. Chem. B* **2005,** *109,* 5895–5902.

[178] Lynden-Bell, R. M.; Del Popolo, M. G.; Youngs, T. G. A.; Kohanoff, J.; Hanke, C. G.; Harper, J. B.; Pinilla, C. C. *Acc. Chem. Res.* **2007,** *40,* 1138–1145.

[179] Bhargava, B. L.; Balasubramanian, S. *J. Phys. Chem. B* **2007,** *111,* 4477–4487.

[180] Zahn, S.; Wendler, K.; Delle Site, L.; Kirchner, B. *Phys. Chem. Chem. Phys.* **2011,** *13,* 15083–15093.

[181] Mallik, B. S.; Siepmann, J. I. *J. Phys. Chem. B* **2010,** *114,* 12577–12584.

[182] Mallik, B. S.; Kuo, I.-F. W.; Fried, L. E.; Siepmann, J. I. *Phys. Chem. Chem. Phys.* **2012,** *14,* 4884–4890.

[183] Bodo, E.; Sferrazza, A.; Caminiti, R.; Mangialardo, S.; Postorino, P. *J. Chem. Phys.* **2013,** *139,* 144309.

[184] Payal, R. S.; Balasubramanian, S. *Phys. Chem. Chem. Phys.* **2014,** *16,* 17458–17465.

[185] Campetella, M.; Bodo, E.; Caminiti, R.; Martino, A.; D'Apuzzo, F.; Lupi, S.; Gontrani, L. *J. Chem. Phys.* **2015,** *142,* 234502.

[186] Mondal, A.; Balasubramanian, S. *J. Phys. Chem. B* **2015,** *119,* 1994–2002.

[187] Firaha, D. S.; Kirchner, B. *J. Chem. Eng. Data* **2014,** *59,* 3098–3104.

[188] Firaha, D. S.; Kirchner, B. *ChemSusChem* **2016,** *9,* 1591–1599.

[189] Yadav, A.; Pandey, S. *J. Chem. Eng. Data* **2014,** *59,* 2221–2229.

[190] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983,** *72,* 926–935.

[191] Nosé, S. *J. Chem. Phys.* **1984,** *81,* 511–519.

[192] Hoover, W. G. *Phys. Rev. A* **1985,** *31,* 1695–1697.

[193] McGrath, M. J.; Kuo, I.-F. W.; Siepmann, J. I. *Phys. Chem. Chem. Phys.* **2011,** *13,* 19943–19950.

[194] Zhao, W.; Leroy, F.; Heggen, B.; Zahn, S.; Kirchner, B.; Balasubramanian, S.; Muller-Plathe, F. *J. Am. Chem. Soc.* **2009,** *131,* 15825–15833.

[195] Wernet, P.; Nordlund, D.; Bergmann, U.; Cavalleri, M.; Odelius, M.; Ogasawara, H.; Näslund, L. Å.; Hirsch, T. K.; Ojamäe, L.; Glatzel, P.; Pettersson, L. G. M.; Nilsson, A. *Science* **2004,** *304,* 995–999.

[196] Brehm, M.; Kirchner, B. *J. Chem. Inf. Model.* **2011,** *51,* 2007–2023.

[197] Lewars, E. G. *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics;* Springer Science & Business Media: Dordrecht, The Netherlands, 2nd ed.; 2011.

[198] Marzari, N.; Vanderbilt, D. *Phys. Rev. B* **1997,** *56,* 12847–12865.

[199] Thomas, M.; Brehm, M.; Fligg, R.; Voehringer, P.; Kirchner, B. *Phys. Chem. Chem. Phys.* **2013,** *15,* 6608–6622.

[200] Liu, H.; Maginn, E.; Visser, A. E.; Bridges, N. J.; Fox, E. B. *Ind. Eng. Chem. Res.* **2012,** *51,* 7242–7254.

[201] Kuo, I.-F. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I. *J. Chem. Theory Comput.* **2006,** *2,* 1274–1281.

[202] D'Agostino, C.; Harris, R. C.; Abbott, A. P.; Gladden, L. F.; Mantle, M. D. *Phys. Chem. Chem. Phys.* **2011,** *13,* 21383–21391.

[203] Sanz-Pérez, E. S.; Murdock, C. R.; Didas, S. A.; Jones, C. W. *Chem. Rev.* **2016,** *116,* 11840–11876.

[204] Abanades, J.; Arias, B.; Lyngfelt, A.; Mattisson, T.; Wiley, D.; Li, H.; Ho, M.; Mangano, E.; Brandani, S. *Int. J. Greenh. Gas Con.* **2015,** *40,* 126–166.

[205] Wilhelm, S.; Bloom, N. *Fuel Process. Technol.* **2000,** *63,* 1–27.

[206] Portney, P. R. *J. Econ. Perspect.* **1990,** *4,* 173–181.

[207] Markovs, J.; Heights, Y.; Cintins, P. E.; Plaines, D. US Pat., 4874525A, 1989.

[208] Li, L.; Li, X.; Lee, J.-Y.; Keener, T. C.; Liu, Z.; Yao, X. *Ind. Eng. Chem. Res.* **2012,** *51,* 9136–9144.

[209] Granite, E. J.; Pennline, H. W.; Hargis, R. A. *Ind. Eng. Chem. Res.* **2000,** *39,* 1020–1029.

[210] Eckersley, N. *Hydrocabon Process.* **2010,** 29–35.

[211] Hiroshi, N.; Toshio, A.; Katsuya, N. US Pat., 4500327A, 1985.

[212] Rogers, R. D.; Holbrey, J.; Rodriguez, H. WO Pat., 116165A2, 2010.

[213] Abai, M.; Atkins, M. P.; Hassan, A.; Holbrey, J. D.; Kuah, Y.; Nockemann, P.; Oliferenko, A. A.; Plechkova, N. V.; Rafeen, S.; Rahman, A. A.; Ramli, R.; Shariff, S. M.; Seddon, K. R.; Srinivasan, G.; Zou, Y. *Dalton Trans.* **2015,** *44,* 8617–8624.

[214] Ji, L.; Thiel, S. W.; Pinto, N. G. *Ind. Eng. Chem. Res.* **2008,** *47,* 8396–8400.

[215] Ji, L.; Thiel, S. W.; Pinto, N. G. *Water, Air, Soil Pollut. Focus* **2007,** *8,* 349–358.

[216] Rogers, R. D.; Holbrey, J. US Pat., 20120090430A1, 2012.

[217] Dai, Y.; van Spronsen, J.; Witkamp, G.-J.; Verpoorte, R.; Choi, Y. H. *Anal. Chim. Acta* **2013,** *766,* 61–68.

[218] Hayyan, A.; Mjalli, F. S.; AlNashef, I. M.; Al-Wahaibi, Y. M.; Al-Wahaibi, T.; Hashim, M. A. *J. Mol. Liq.* **2013,** *178,* 137–141.

[219] Garcia, G.; Aparicio, S.; Ullah, R.; Atilhan, M. *Energy Fuels* **2015,** *29,* 2616–2644.

[220] Mahto, A.; Mondal, D.; Polisetti, V.; Bhatt, J.; M. R, N.; Prasad, K.; Nataraj, S. K. *Ind. Eng. Chem. Res.* **2017,** *56,* 14623–14632.

[221] AlOmar, M. K.; Alsaadi, M. A.; Jassam, T. M.; Akib, S.; Hashim, M. A. *J. Colloid Interface Sci.* **2017,** *497,* 413–421.

[222] Smith, J. G. *Organic Chemistry;* McGraw-Hill: New York, 2nd ed.; 2008.

[223] Mancini, M. V.; Spreti, N.; Di Profio, P.; Germani, R. *Sep. Purif. Technol.* **2013,** *116,* 294–299.

[224] Cheng, G.; Bai, B.; Zhang, Q.; Cai, M. *J. Hazard. Mater.* **2014,** *280,* 767773.

[225] ShamsiJazeyi, H.; Kaghazchi, T. *J. Ind. Eng. Chem.* **2010,** *16,* 852–858.

[226] Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard III, W. A. and Skiff, W. M. *J. Am. Chem. Soc.* **1992,** *114,* 10024–10035.

[227] Hartwingsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998,** *58,* 3641–3662.

[228] Grossfield, A. WHAM: the weighted histogram analysis method, version 2.0.9, 2016.

[229] Earnwood, N. N.; Greenshaw, A. *Chemistry of the Elements;* Elsevier B.V.: Oxford, 2nd ed.; 1997.

[230] Fetisov, E. O.; Siepmann, J. I. *J. Phys. Chem. B* **2016,** *120,* 1972–1978.

[231] Gu, T.; Whitesell, J. K.; Fox, M. A. *J. Org. Chem.* **2004,** *69,* 4075–4080.

[232] Ansar, S. M.; Gadogbe, M.; Siriwardana, K.; Howe, J. Y.; Dogel, S.; Hosseinkhannazer, H.; Collier, W. E.; Rodriguez, J.; Zou, S.; Zhang, D. *J. Phys. Chem. C* **2014,** *118,* 24925–24934.

[233] Salvio, R.; Cincotti, A. *RSC Adv.* **2014,** *4,* 28678–28682.

[234] De, M.; Ghosh, P. S.; Rotello, V. M. *Adv. Mater.* **2015,** *20,* 4225–4241.

[235] Mrksich, M.; Sigal, G. B.; Whitesides, G. M. *Langmuir* **1995,** *11,* 4383–4385.

[236] Ma, F.; Lennox, R. B. *Langmuir* **2000,** *16,* 6188–6190.

[237] Shon, Y.-S.; Lee, S.; Perry, S. S.; Lee, T. R. *J. Am. Chem. Soc.* **2000,** *122,* 1278–1281.

[238] Jackson, A. M.; Myerson, J. W.; Stellacci, F. *Nat. Mater.* **2004,** *3,* 330–336.

[239] Cesbron, Y.; Shaw, C. P.; Birchall, J. P.; Free, P.; Lévy, R. *Small* **2012,** *8,* 3714–3719.

[240] Stirling, J.; Lekkas, I.; Sweetman, A.; Djuranovic, P.; Guo, Q.; Pauw, B.; Granwehr, J.; Lévy, R.; Moriarty, P. *PLoS ONE* **2014,** *9,* e108482.

[241] Singh, C.; Ghorai, P. K.; Horsch, M. A.; Jackson, A. M.; Larson, R. G.; Stellacci, F.; Glotzer, S. C. *Phys. Rev. Lett.* **2007,** *99,* 226106.

[242] Ghorai, P. K.; Glotzer, S. C. *J. Phys. Chem. C* **2010,** *114,* 19182–19187.

[243] Devi, J. M. *J. Mol. Model.* **2015,** *21,* 149.

[244] Velachi, V.; Bhandary, D.; Singh, J. K.; Cordeiro, M. N. D. S. *J. Phys. Chem. C* **2015,** *119,* 3199–3209.

[245] Van Lehn, R. C.; Alexander-Katz, A. *J. Phys. Chem. C* **2013,** *117,* 20104–20115.

[246] Bain, C. D.; Evall, J.; Whitesides, G. M. *J. Am. Chem. Soc.* **1989,** *111,* 7155–7164.

[247] Bain, C. D.; Whitesides, G. M. *J. Am. Chem. Soc.* **1989,** *111,* 7164–7175.

[248] Siepmann, J. I.; McDonald, I. R. *Mol. Phys.* **1992,** *75,* 255–259.

[249] Sun, L.; Siepmann, J. I.; Schure, M. R. *J. Phys. Chem. B* **2006,** *110,* 10519–10525.

[250] Siepmann, J. I.; Frenkel, D. *Mol. Phys.* **1992,** *75,* 59–70.

[251] Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1999,** *103,* 4508–4517.

[252] Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998,** *5647,* 2569–2577.

[253] Lubna, N.; Kamath, G.; Potoff, J. J.; Rai, N.; Siepmann, J. I. *J. Phys. Chem. B* **2005,** *109,* 24100–24107.

[254] Siepmann, J. I.; McDonald, I. R. *Mol. Phys.* **1993,** *79,* 457–473.

[255] Maitland, G. C.; Rigby, M.; Smith, E. B.; Wakeham, A. *Intermolecular Forces: Their Origin and Determination;* Pergamon Press: Oxford, 1987.

[256] Zhao, X.; Leng, Y.; Cummings, P. T. *Langmuir* **2006,** *22,* 4116–4124.

[257] Nelson, D. R. *Nano Lett.* **2002,** *2,* 1125–1129.

[258] Ghorai, P. K.; Glotzer, S. C. *J. Phys. Chem. C* **2007,** *111,* 15857–15862.